THE FUNCTION OF ELABORATIVE RESPONSES IN THE PROCESSING OF A PERSUASIVE COMMUNICATION

BY

JOHN W. VANN

A DISSERTATION PRESENTED TO THE GRADUATE COUNCIL
OF THE UNIVERSITY OF FLORIDA IN
PARTIAL FULFILLMENT OF THE REQUIREMENTS
FOR THE DEGREE OF DOCTOR OF PHILOSOPHY

UNIVERSITY OF FLORIDA

1980

To my parents

Flora H. Vann and Russell F. Vann

ACKNOWLEDGEMENTS

I want to thank the experimental team who worked so hard in assembling the paraphernalia used in the experimental sessions. Terry Armstrong, Meegan Carol, Peter Gold, Sally Krusing and Alan Moore not only worked hard physically, but also carried out the experimental procedures smoothly and tactfully handled all subject problems which arose.

Mary B. Faricy and Barbara S. Probert were also very helpful, not only through their role of securing subjects as principals of Central Florida Research, but also through their help in the preparation for and conduct of the experimental sessions.

Thanks are also due my committee. Professors Barry R. Schlenker and William L. Wilkie provided useful observations and discussion regarding the conceptualizations which were used in this dissertation. My committee chairman, Professor Joel B. Cohen, provided useful, critical reading of and feedback on all parts of the dissertation. In addition, as director of the Center for Consumer Research, he along with Professor Wilkie, the principal investigator of Experiment 1 described in Chapter 5, provided the opportunity and the subjects for my experiment.

A very special thanks is due my wife, Carolyn. She labored many, sometimes very late, hours typing the dissertation in its various forms. As an intelligent and thoughtful reader, she helped me to purge many of the logical and grammatical errors which I created. Most importantly, she believed in me. Her support and encouragement kept me going throughout this endeavor.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

"...only when science has reached the stage where it can say 'there are no minds' will it have accounted adequately for mind."

Brodbeck 1966, pp. 57-58

"As the well known saying goes: psychology first lost its soul, later its consciousness, and seems now in danger of losing its mind altogether."

Feigl 1959, p. 121

"For isn't it stupid to hoot at people whose only fault is that their ancestors bequeathed them a shabby mesh of associative nerve fibers."

Pitkin 1932, p. 8

Abstract of Dissertation Presented to the Graduate Council
of the University of Florida in Partial Fulfillment of the
Requirements for the Degree of Doctor of Philosophy

THE FUNCTION OF ELABORATIVE RESPONSES IN THE
PROCESSING OF A PERSUASIVE COMMUNICATION

by

John W. Vann

August 1980

Chairman: Joel B. Cohen
Major Department: Marketing

The purpose of this dissertation is to further understanding of the role of elaborative responses in the processing of persuasive communications. A theoretical background of information processing is provided with special attention given to the nature and role of sub-vocal responses. A review of the persuasion literature which presumes an intervening role of elaborative responses is presented with a focus upon those studies which examined the effects of distraction on persuasion. Methodological guidance is provided for doing experimentation in the realm of elaborative responses.

The experimental approach consisted of experimentally manipulating the assailability of the pre-message information (high and low), the credibility of the message source (high and low) and the discrepancy between the pre-message position and the position expressed in the message (low, moderate and high). Verbal protocols of reported elaboration were elicited along with measures of message acceptance. A manipulation check revealed that the manipulations of assailability and source

credibility were deficient. Consequently, causal conclusions regarding the effects of these variables may not be made. Increasing discrepancy from low to moderate increased those negative protocol statements as a proportion of all protocol statements. High discrepancy resulted in a significant shift in pre-message position when compared with a no-message control. Correlational analysis indicated that coded protocols were significantly related to other criterion measures, with the proportion of negative statements being the best predictor.

An internal analysis examined the relationship between manipulation measures and criterion measures. This correlational analysis indicated that the criterion measures which were relevant to the message were related to source credibility, while those relevant to the background information were related to assailability. However, the expected cross-relationships were not observed. Suggestions for further research in the area are presented.

CHAPTER 1

INTRODUCTION

Probably everyone has had the subjective experience of images being aroused by a written or spoken communication or of silent dialogs carried on with the communicator. Sometimes these subvocal responses embellish the message--adding associated ideas or logical extension. Sometimes they may refute a part of the message or deprecate the communicator. There may be some responses of this nature that a person is not even aware of or cannot recall immediately after they have passed fleetingly through the person's mind. All of these responses are called elaborative responses and are the focus of the present study.

Elaborative responses are important because there is evidence that they play an essential role in integrating message content into the message recipient's cognitive structure. Through this integrative function, elaborative responses influence the message-topic position of the message recipient as reflected by his or her post-message cognitive structure. This impact on post-message cognitive structure suggests that an understanding of the role of elaborative responses has direct relevance to a more complete understanding of message-induced persuasion. (In this case persuasion would be considered to be a shift in the message recipient's position such that the position held after the communicative episode is closer to that espoused by the communicator than was the pre-communication position.)

The purpose of this study is to contribute to the understanding of the role of elaborative responses in the processing of persuasive communications. However, the focus here is on cognitive changes which may

result from a communicative episode, not on behavioral manifestations of those changes.

A thorough understanding of the role of elaborative responses in message processing should include knowledge of the antecedents as well as the consequents of elaborative responses. A model which portrays the integrative role of elaborative responses in message processing is presented in Figure 1. This model also illustrates antecedent variables which may be expected to influence elaborative responses, and, subsequently, the post-message cognitive structure.

The central role of elaboration in this model sets it apart from models of persuasion which view the message recipient as a passive participant in the persuasive process--who needs only to receive and comprehend the message to be persuaded. Instead, this model views the recipient as an active, subvocal participant in a dialog with the sender. The cognitive residue of the communication, rather than being seen as a copy of the message as it was sent, is seen as a complex mosaic resulting from the interaction of message content and the sometimes-verbal elaborative responses of the recipient. This complex, post-message, cognitive representation of the communicative episode will reflect the influence of the variables which impinge upon elaboration in the model. Because of this influence, there are important implications of the model for anyone concerned with persuasive communications. For example, the model suggests that marketers should concern themselves with <u>message variables</u> such as the number and order of arguments used in a message.

Another important influence on elaboration, as presented in the model, is the sum of all variables which provide the <u>context</u> within

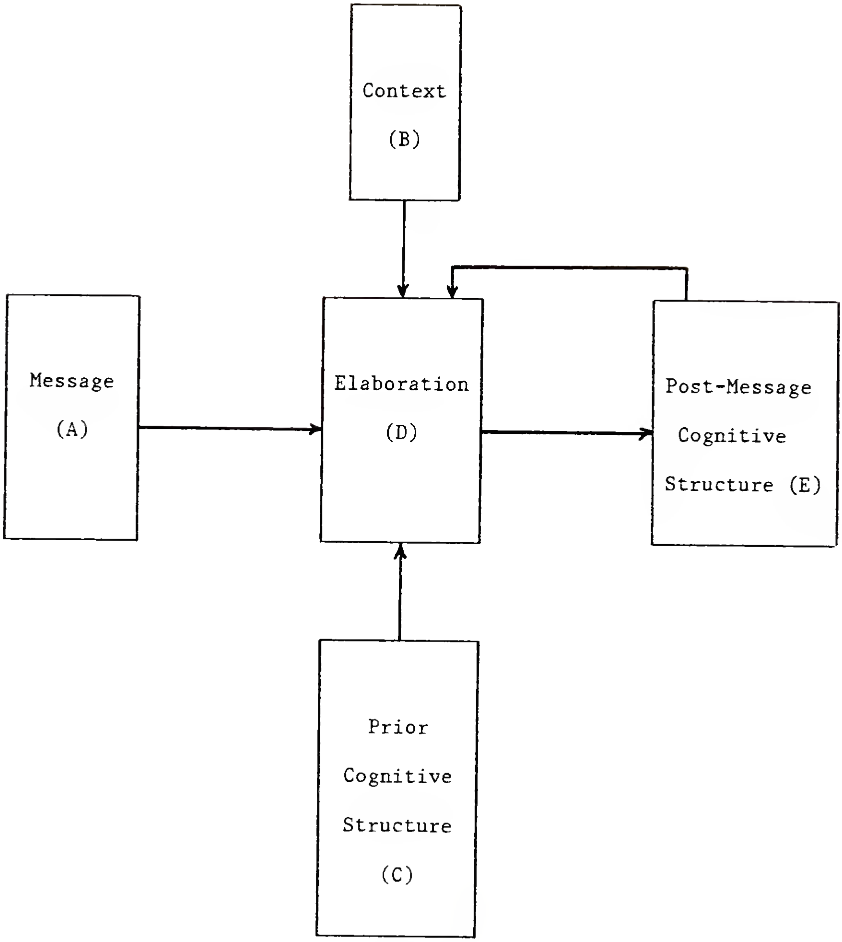


Figure 1. An Elaboration Model of Message Processing

which the message is presented. The message medium, source characteristics, as well as concurrent surrounding stimuli are all seen as contextual variables which influence message elaboration. The impact of source credibility can be especially important to advertisers, who, because of their vested interest in the persuasive outcome, may suffer from chronic low credibility as perceived by the message recipients. A low-credibility source would be expected to engender more negative responses, and, hence more negative post-message cognitive structure, during elaboration than would a high-credibility source.

An important consideration to both marketers and those involved in setting public policy is the limitation in time and processing capacity which must be allocated among competing cognitive tasks. Thus, an important aspect of the <u>context</u> within which the message episode occurs is the presence or absence of distracting stimuli which compete with the message for processing resources. There is evidence that, under some circumstances, distracted message recipients are more easily persuaded. Public policy makers may be concerned that the use of distractors in advertising may take unfair advantage of the consumer's processing limitations. The literature relevant to the impact of distraction on elaboration and persuasion is reviewed in Chapter 3.

The <u>prior cognitive structure</u> of the message recipient is also seen as an important influence upon elaboration. The pre-message position of the recipient is shown as influencing elaboration and thereby the recipient's post-message position. This suggests that a marketer should consider the beliefs and values of his target market prior to message preparation. The discrepancy between the position expressed in the message and the positions held by the recipients would be shown in the

elaboration model as an interaction between message and prior-cognitive-structure variables. Because of the influence of discrepancy on elaboration marketers may find it useful to prepare different messages for market segments which hold divergent message-relevant positions.

The present study represents an attempt to manipulate some of the antecedent variables which were hypothesized to influence the elaborative process. However, methodological problems arose in the operationalization of two of the antecedent variables involved: the credibility of the message source (a contextual variable) and the assailabilty of the information upon which the respondents based their pre-message position (a prior cognitive structure variable). Manipulation check measures taken during the experimental sessions revealed that the manipulation of assailability was completely ineffective. The manipulation of source credibility was deficient in that only one dimension of a multidimensional variable was affected. Consequently none of the questions regarding the influence of these two varaibles on elaboration or on persuasion may be answered in a causal sense. Because of the methodological problems in the experiment, it seemed appropriate to play down the experimental results and to focus instead on the theoretical and methodological issues associated with the research problem.

This study will be presented in three parts. Part one (Chapters 2 and 3) presents a review of the literature relevant to the theoretical background of information processing and then examines studies of the relationships portrayed in the elaboration model. Chapter 2 provides the theoretical background for the study of message processing. This background includes discussions of different theoretical positions

regarding both cognitive structures and cognitive processes. Methodological issues associated with the study of cognitive processes are also discussed.

Because of the central role in message processing proposed for elaborative responses in the model, the literature relevant to the existence of the functions of and the limitations on elaborative responses are examined extensively. Elaborative responses are presented as a major determinant of the representation of message content as it is stored in memory. The extent of elaboration is presented as a key factor influencing the duration of time over which the message content may be recalled. Limitations of processing capacity are also presented and the implications of restricting elaboration are discussed. The chapter ends with a consideration of possible ways of monitoring the occurrence and content of the elaborative process.

Chapter 3 presents a review of studies which have examined the relationships portrayed in the elaboration model. The greater part of this chapter deals with studies of the effects of distraction (a contextual variable). These studies involve the use of distractors which may limit the processing capacity that may be allocated to message-relevant elaborative responses. Other studies reviewed in this chapter include some which have examined the effects of message, prior-cognitive-structure and contextual variables (other than distraction) on reported elaboration. Some other studies reviewed have examined the relationship between reported elaboration and measures of message acceptance. The large literature which has examined verbal protocols as measures of elaboration is also examined in this chapter. The findings of all of these studies are examined to see if they corroborate the relationships represented in the model.

Some of the variables represented in the elaboration model involve variables which may be manipulated and observed directly (such as the number of arguments in a message). Others require some response by an experimental subject for their manipulation (such as the perceived credibility of a source) and may not be directly observed. Part two (Chapters 4 and 5) examines the methodological issues surrounding the manipulation of experimental variables and attempts to determine the causal nature of relationships such as those depicted in the elaboration model.

Chapter 4 examines some of the methodological problems which may arise in trying to manipulate unobservable variables such as those employed in the present study. This chapter also presents other sources of variance which may intrude upon attempts to establish the causal role of experimental variables (such as the antecedent variables in the elaboration model) as well as methods for controlling these sources of variance. Chapter 5 presents the hypotheses and methods employed in the present study.

Part three (Chapter 6) will present the results of the study as conducted. None of the hypotheses involving the credibility or assailability variables could be tested because of the faulty manipulation of those two variables. However, the results of the third predictor variable (discrepancy) which was manipulated orthogonally to credibility and assailability will be presented. Two additional analyses will be presented. The first will examine correlations between various criterion measures and the second will involve an internal analysis.

This internal analysis will examine the correlations between the <u>measured</u> indicants of perceived source credibility and perceived

information assailability with the criterion variables. "Such an analysis... cannot prove a causal relationship, any more than a correlational study can outside the context of an experiment, but it can be extremely useful as a source of information for guiding future experimentation" (Carlsmith, Ellsworth and Aronson 1976, p. 37). The chapter will end with suggestions for future research into the role of elaboration in the communicative process.

CHAPTER 2

THEORETICAL BACKGROUND ON COGNITIVE STRUCTURES AND PROCESSES

There have been many persuasion studies, both from a psychological and a marketing viewpoint which have presumed a mediating role of subvocal (i.e., elaborative) responses in the persuasion process (e.g., Festinger and Maccoby 1964; Brock 1967; Greenwald 1968; Wright 1973a, 1974b, 1975). Some of these studies have even purported to measure these subvocal responses (e.g., Brock 1967; Greenwald 1968; Wright 1973a, 1974b, 1975). These studies will be the focus of Chapter 3. However, these studies are an applied part of a broader field of study--the study of cognitive structures and processes. There are many substantive issues which have been examined in the history of this field and these issues should be considered in the evaluation of the persuasion studies reviewed in Chapter 3. The purpose of this chapter (2) is to present the theoretical and methodological issues associated with the study of cognitive structures and processes. This material will then comprise the conceptual framework within which studies which have examined the relationships in the elaboration model may be evaluated.

The present chapter is concerned with how people extract information from oral and written communications, how a person may transform this verbal material and how it would be stored in and retrieved from memory. In short, it is concerned with cognitive <u>structures</u> and <u>processes</u>.

The discussions herein relate to word recognition, memory structure, sentence comprehension, text processing, sub-vocal processes, the resolution of message-belief discrepancies, message recall, and processing limitations. The role of subvocal processes (i.e., elaborative responses) is presented as lying primarily in the stage of transformation of verbal material with implications for persuasion, storage and retrieval. The chapter ends with a consideration of how elaborative processes may be monitored.

Word Recognition

Word recognition, whether from spoken or written material, requires that a person compare the word or symbols of each word with some standard in memory so that the word meaning can be established (Norman 1976). The set against which the incoming words are compared has been variously called a mental thesaurus (Cofer 1976; Tulving 1972), a subjective lexicon (Conrad 1974), a lexicon (Kintsch 1975), and a semantic memory (Norman 1976; Tulving 1972).

On occasion, words must be referenced against memory for specific episodes (episodic memory) in which incidental use of the words is recalled (LaBerge and Samuels 1974). Norman and Bobrow (1976) suggest that presentation of a word can result in the activation of a schema. Schemata are "...semi-independent procedures that analyze the data sent to them and return results to some common data pool..." (1976, p. 119) (see Figure 2).

These schemata are comparable to Schneider and Shiffrin's (1977) "nodes" which "may consist of a complete set of informational elements, including associative connections, programs for responses or actions, and directions for other types of information processing" (1977, p. 2).

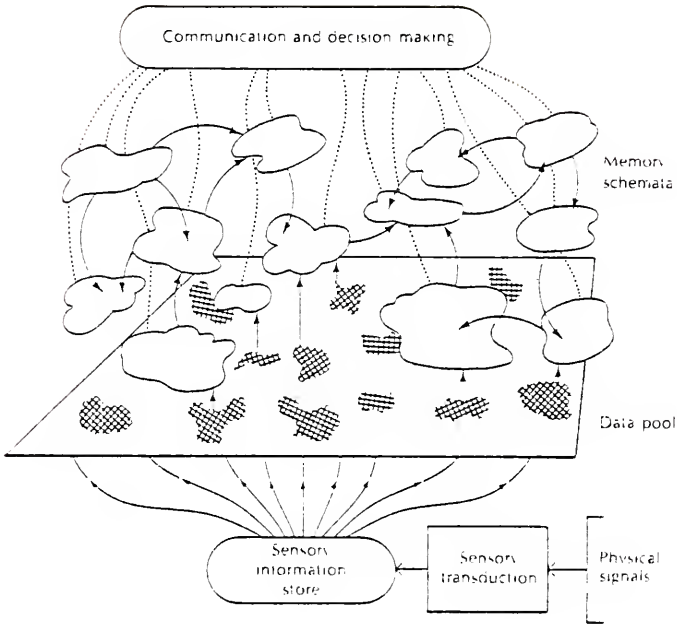


Figure 2. "The memory schemata view of the human information-processing system. Incoming data and higher-order conceptual structures all operate together to activate memory schemata. Short-term memory consists of those schemata that are undergoing active processing. There is no set of sequential stages: the limits on processing capability are set by the total amount of processing resources available to the system." (From Norman and Bobrow, 1976, p. 118).

Semantic Memory:

The semantic memory to which messages are referred can be considered as a set of interrelated descriptions (Norman and Bobrow 1976). These descriptions define concepts by linking concepts with their properties. Concepts become linked with their properties through repeated exposure to occurrences of the concept and properties accompanying each exposure to the concept. In this sense semantic memory is derived from episodic memory and is, therefore, ultimately based upon sensory images (Norman 1976, p. 189; Lindsay and Norman 1977, p. 399). However, this means that our representation of reality may be faulty. "Not only shall I insist that subjective experience cannot provide the true foundation of our knowledge, but I shall insist that our knowledge has no foundations in the traditional sense.... I simply mean that there is no such thing as an indubitable foundation, on which knowledge of any sort can rest" (Aune 1967, p. 38).

The concomitant occurrence of some concepts and properties is random. These properties are eliminated over time as the concept description is refined. Other properties occur simultaneously with the concept only under certain circumstances--thus the term "context-dependent descriptions" (Norman and Bobrow 1976, p. 127). Subsequent events or propositions are identified and assigned to the appropriate conceptual categories (i.e., perceived) according to which properties are present. Assignment rules include: a) the requisite properties; b) the way in which properties must be combined; c) the weights assigned to the various properties; and d) the acceptable range of dimensional values of the properties (Bruner 1957, pp. 131-132).

When a context-dependent description is so unique as to include linkages to a specific time and place, then the representation of that concept is considered the record of an event and a part of episodic, rather than semantic memory (Tulving 1972). Lindsay and Norman (1977, p. 397) list action, agent, conditionality (e.g., "only if" or "because"), instrument, location, object, recipient, time, and truth as parts of an event. As instances of a concept occur in more and more contexts, it becomes relatively context-free (Norman and Bobrow 1976; Ausubel 1963, p. 218).

The descriptions can include linkages between concepts and properties implying "...something positive or negative, something strong or weak, something active or passive. There exists a large literature concerned with this problem using the semantic differential as an instrument (Osgood, Suci, and Tannenbaum 1957)..." (Kintsch 1974, p. 24).

The linkages between a concept and its properties may reflect some uncertainity. A method of expressing this uncertainty for noun-type concepts is included in several attitude models (Fishbein and Ajzen 1975; Ahtola 1975). In these models, the linkages are called beliefs and are expressed as subjective probabilities. (See Jones and Gerard 1967, p. 160, for an equivalent, syllogistic treatment of attitude as a derivative of this linkage).

Inferred properties:

The property-concept linkages need not be present in memory to be used by a person. They may be potentially derivable from existing linkages through inference rules. Descriptions of more general concepts are assumed to apply also to those concepts that are subordinate, unless

otherwise specified (Lindsay and Norman 1977, pp. 386-388). The inference process can be represented as a syllogistic one: All _A_s have _B_s; a _C_ is an _A_; therefore, _C_s must have _B_s. Of course, this process is dependent upon the _is a_ linkage as part of the description of a _C_ (Jones and Gerard 1967, p. 160; Lindsay and Norman 1977, p. 388). Concepts can also be linked to higher order concepts through a similar process: All _A_s are _B_s; all _B_s are _C_s; therefore, all _A_s are _C_s (Kintsch 1974, p. 26).

Because of the extensive implications derivable from general concepts to subordinate concepts, changes in schemata (descriptions) for general concepts will result in greater overall change in the cognitive structure (extending to all subordinate concepts) than will changes in descriptions that are limited to a subordinate concept. Changes that are implied in one part of the cognitive structure by changes in another part may not take place immediately. This reflects a characteristic called "cognitive inertia" (McGuire 1968, p. 153). However, the very interrelatedness of concepts in semantic memory protects descriptions from change induced by incoming information (Tulving 1972, p. 391).

Consequences as properties. Some descriptions refer to actions. Lexically, these would be expressed as verbs. The description for the verb-type concept _HIT_ includes "...slots for an _agent_, an _object_, and an _instrument_..." as well as possible consequences of the actions (Kintsch 1977, p.343). Consequences of actions are also concepts.

One classification scheme for consequences has been developed by Rokeach (1968a, 1968b, 1968c). These properties of actions are divided into terminal values (end states) and instrumental values (appropriateness of a mode of behavior). Rokeach (1968a) considers the degree

of interconnectedness of a value in the cognitive structure is a measure of its centrality. Those values that have more interconnections within the structure are considered more central. In this way, centrality is comparable to generality as more general concepts are directly connected to all of the subordinate concepts, while each subordinate concept is only directly connected to the general concept.

Rokeach equates centrality with importance. "The parts are conceived to be arranged along a central-peripheral dimension wherein the more central parts are more salient or important, more resistant to change, and if changed, exert relatively greater effects on other parts" (Rokeach 1968a, p. 117).

Rosenberg (1956) developed an attitude model that related the linkage between actions and consequences in a probabilistic fashion similar to that linking noun-type concepts and properties in the models discussed earlier. Wright (1973b) points out that attitudes toward noun-type concepts ultimately rest on the consequences linked to the properties that are linked with the noun-type concept. This chaining is an example of what Jones and Gerard called a vertical attitude structure (1967, p. 184). While such chaining would maintain internal cognitive consistency and may describe an inferential process, evaluative linkages need not be formed in this manner.

Schema Activation:

A particular schema is activated when cues are present which match the "context-dependent description" appropriate to that schema (Bobrow and Norman 1975, p. 133; Norman and Bobrow 1976, p. 127).

If the physical and contextual elements that are present do not uniquely describe a particular memory schema for a word, then ambiguity

will exist and several schemata may be activated (Bobrow and Norman 1975). These schemata can be activated either by external stimulation or by a person's expectations. The activation is considered data-driven to the extent that the schemata are activated by external stimulation. Activation which is internal is considered conceptually-driven. "Conceptually driven processing tends to be top-down, driven by motives and goals, and fitting input to expectations; event driven processing tends to be bottom-up, finding structures in which to embed the input" (Bobrow and Norman 1975, p. 140). Communications involve both data-driven and conceptuallly-driven processing (as will be discussed further in the section on text processing).

A similar "logogen" model was proposed by Morton (1969; see also Keele 1973; and Posner and Warren 1972). Morton called a logogen a "device" which collects information in the form of auditory, visual, and semantic (contextual) attributes. When the number of attributes exceeds a threshold level, then the logogen makes its output available. Keele suggests that this output can be "...the name of the stimulus, the meaning of the stimulus, or an appropriate response to the stimulus" (Keele 1973, p. 88). All other logogens for which the attributes are appropriate accumulate those attributes, but unless the threshold is reached, no output is forthcoming. In all cases the context of the communication influences what word is perceived. The context includes grammatical, situational, and cultural cues as to what words will be a part of the communication.

Another, similar model has been proposed by Hayes-Roth (1977). Her model proposes that our cognitive structure is composed of elementary memory units or <u>cogits</u>, "...the smallest information structure

perceptually or cognitively delineated. Each cogit is identified with a discrete memory representation" (p. 261). These cogits are assembled into larger configurations through association. These configurations can become so strongly associated that they become "unitized." "A unitized assembly may be reconstructively activated in an all-or-none fashion by a stimulus that includes a subset of the information represented in the assembly or by activation of the internal representation of such a stimulus" (p. 262). This activation is essentially comparable to schema or logogen activation.

Contextual stimulation of a category in memory makes that category more accessible because of expectations regarding "...likely transitions and contingencies of the environment" (Bruner 1957, p. 135). Bruner suggests that the more category accessibility is increased, the less additional input is required for categorization into a given category (as in the logogen model). "...The wider the range of input characteristics that will be 'accepted' as fitting the category in question [and]...the more likely that categories that provide a better or equally good fit for the input will be masked" (Bruner 1957, pp. 129-130).

Once a particular memory referent has been activated, it can in turn activate others (Norman and Bobrow 1976). "A word is more than an arbitrary written or spoken sign, it is all that it carries in associations as well" (Cherry 1966, p. 72). These associations, in turn, can generate expectations of subsequent words, as well as enriching the meaning of the presented word.

An important property of the initial word identification process is its automaticity for familiar words (Deutsch and Deutsch 1963; LaBerge and Samuels 1974). This property has also been called "effortless

retrieval" (Posner 1973, p. 40). It is significant because it does not interfere with other cognitive processes. This property will be covered in more depth in the discussion of interference and distraction.

Sentence Comprehension

The Use of Redundancy in Sentence Comprehension:

Words that have been identified can be retained for a short period of time in memory without rehearsing them (10 - 20 seconds). This is long enough for an entire phrase to be completed. While reading or listening to prose material, it is not necessary that every word be processed since the receiver uses various strategies that depend on the semantic constraints that govern word possibilities and on redundancies that exist in spoken and written prose (Kintsch 1977). "'Redundancy' means additional signs or rules which guard against misinterpretation..." (Cherry 1966, p. 33). "The redundancy contained in an average sentence...provides an important safety factor in the processing of information. We can miss one part of a message but still retrieve the whole because some of the message could have been reliably predicted from the rest of it" (Jones and Gerard 1967, p. 135; see also Bruner 1957, p. 132).

Syntatic Aids to Sentence Comprehension:

Some other strategies are based on frequent patterns that occur in English. One such strategy is assuming that the first noun-verb-noun phrase to occur is the main clause of the sentence, unless the phrase includes "...a cue word such as _that_, _who_, _whom_, or _which_..." to indicate that the main clause is to follow (Kintsch 1977, p. 313). Another strategy is assuming that the order in which events have

occurred is the same as the order of presentation in a sentence (Kintsch 1977, p. 314).

Grammar is another aid in prose comprehension which helps readers or listeners anticipate the order of occurrence of parts of speech, even in the nonsense sentence: "The ventious crapests pounted raditally" (Cherry 1966, p. 121).

Linguistic Relativity and Determinism:

Language has another effect on the processing of verbal messages. It limits the content of messages by the constraints it places on what can be said in a particular language and culture. The language restricts the categories used in describing the world, as well as how these categories are subdivided. As Cherry (1966, p. 73) states, language determines what people "...are free and able to think." This principle is called linguistic relativity and determinism, or sometimes the "'Whorfian hypothesis'"(Slobin 1974, p. 120). A more widely accepted, weaker form of the hypothesis suggests that language primarily affects our epistemology by predisposing us to respond in certain ways, rather than rigidly determining our responses. "One is not fully a prisoner of one's language; it is just a guide to thought and other sorts of behavior" (Slobin 1974, p. 122). "...We tend to say things which can be fairly conveniently encoded and we frequently assimilate experience to the categories of the linguistic code" (Slobin 1974, p. 126).

Text Processing

Text processing involves the extraction of the overall meaning expressed by a series of interrelated sentences. However, any

explication of this process requires a conceptual description of textual meaning.

Meaning Expressed in a Text:

Kintsch (1977) has also developed a representational system for prose. The semantic structure of prose is represented as a text base. "The text base consists of a sequence of propostions; propositions in turn are composed of concepts. Each proposition consists of one relational term and one or more arguments [all of which are concepts] (Kintsch 1977, p. 342). An example of a text base with four propositions would be (Kintsch 1977, p. 343):

(1) (HIT, GEORGE, JOHN)

(2) (APOLOGIZE, GEORGE, (HIT, GEORGE, JOHN))

(3) (VICIOUS, (HIT, GEORGE, JOHN))

(4) (TALL, JOHN, GEORGE)

Proposition (1) could be expressed by the words George hit John. The text could be: Even though John is taller than George, George hit John viciously. Later he apologized. Of course, there are many other combinations of sentences that could express the same underlying propositions.

Internal coherence within a text base is achieved through the repetition of the same argument in multiple propositions as with the repeated occurrence of GEORGE in the example or through the embedding of one proposition as the argument of another, as with the embedding of (HIT, GEORGE, JOHN) in proposition (3).

The advantage of using the text base system of specifying the underlying propositions of a text is twofold: first it provides a means of specifying the meaning that a communicator intends to transmit. This

means that researchers are freed from using sentences as a measure of content in an intended message. This is especially important since many times part of the content in the text base is not included in the text, but must be inferred by the reader or listener (Kintsch 1977, p. 359). Secondly, the resultant text base of the listener or reader can be compared with that intended by the communicator as a means of measuring the accuracy of the communicative process. In most situations, the goal of the reader or listener is to understand the gist of the text as represented by the text base, rather than to retain the surface structure of the message. For example, in listening "...while we generally remember quite well what we have just heard, we often cannot repeat it in the same words in which it was given. Apparently we quickly unravel the meaning and forget the syntax" (Slobin 1974, p. 31).

Using disconnected sentences, Begg and Wickelgren (1974) found that semantic information is both better learned initially and is lost at a slower rate than is lexical/syntactic information. "Sentences are the vehicles by means of which information is transmitted. The vehicle itself is less important than what it contains" (Kintsch 1977, p. 333). This suggests that any insistence on verbatim recall as evidence of communicative success could result in understating the degree to which the intended text base has been received by the reader or listener.

There are also cultural influences on text processing. Not only does culture influence comprehension through the effects of linguistic relativity and determinism, but also through story schemata that are shared within the culture.

Many non-conversational communications such as stories, folk tales, fables, and scientific articles tend to follow accepted, consistent

patterns in a given culture. These patterns provide readers or listeners with an outline which guides comprehension as appropriate information is gleaned from the text (Kintsch 1977, p. 379). These patterns have been formalized by several writers. Rumelhart (1975) provides a story grammar in his discussion of a schema for stories. This grammar can be used to represent both the semantic and syntactic structure of stories.

The Role of Rehearsal in Text Processing:

Figure 3 illustrates the differing extent to which text material can be processed. Processing may be limited to a very elementary level. In reading, the "...meanings of familiar words and word groups may be activated automatically, leaving attention free to wander to other matters" (LaBerge and Samuels 1974, p. 320). A similar phenomenon is found in listening. Often a preoccupied person may hear what someone has said and not realize that the message has been processed. Yet upon asking the speaker to repeat what he said, the "listener" realizes that he knows what was said after all and repetition isn't necessary. This is called the "Oh, Yes" phenomenon (Glucksberg and Cowen 1970, p. 150). Messages processed to this elementary level are lost quickly, unless further processing is performed or unless they are rehearsed by simply repeating them to oneself-- maintenance rehearsal (Craik and Watkins 1973; Craik and Lockhart 1972; Postman 1975). "...maintenance rehearsal is described as a 'telephone strategy', one like the operations performed to keep a new telephone number in mind from the time one gets it until it can be dialed on the telephone" (Norman 1976, p. 119). If the person is sufficiently motivated, the attentional demands allow it, and the material is amenable, more extensive processing may take place.

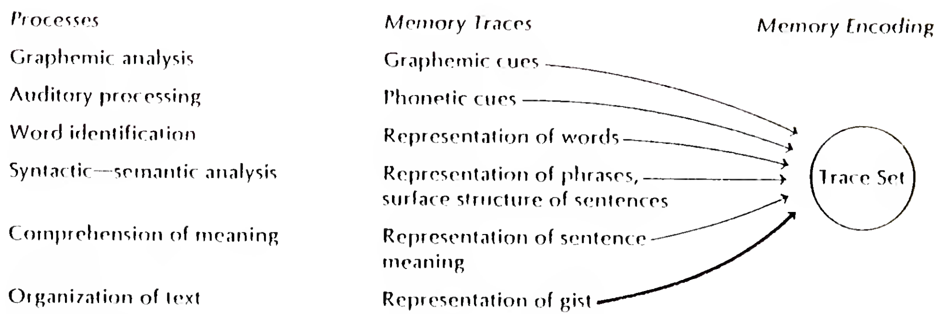


Figure 3. "An Outline of Stages of Comprehension in Reading, the Memory Traces Arising from these Processes, and Their Selection for Encoding in Memory. The Thickness of the Arrows Indicates the Likelihood that Particular Types of Memory Traces Will be Encoded in Memory for Long Term Retention under Standard Reading Conditions. (after Kintsch, 1975)" (from Kintsch 1977, p. 329).

Craik and Lockhart (1972, p. 675) propose a continuum of processing from "...a series of sensory stages to levels associated with matching or pattern recognition and finally to semantic-associative stages of stimulus enrichment." (See also Postman 1975, p. 303; Norman and Bobrow 1976). More extensive processing is called elaborative rehearsal (Craik and Watkins 1973; Norman 1976, p. 127) or constructive rehearsal (Postman 1975, p. 301). This can occur at the word level or at the sentence or text level. "...after a word is recognized, it may trigger associations, images or stories on the basis of the subject's past experience with the word" (Craik and Lockhart 1972, p. 675). LaBerge and Samuels (1974) suggest that the same process can occur in reading prose.

Message Elaboration:

Elaboration is a process of relating the incoming message to existing knowledge in semantic memory and to event-based memory. Event-based memory is called episodic memory. "The remembered episodes...are autobiographical events, describable in terms of their perceptible dimensions or attributes and in terms of their temporal-spatial relations to other such events" (Tulving 1972, p. 387). "Most, if not all, episodic memory claims a person makes can be translated into the form: 'I did such and such, in such and such a place, at such and such a time'" (Tulving 1972, p. 389). Postman (1975, p. 325) points out that we can recall many aspects of an event. Not only can we recall what, when, where, and how, but also how frequently an event has occurred.

The role of implications in the elaboration process:

Elaboration also requires the assumption of certain facts implied in the text. Lindsay and Norman (1977, p. 485) illustrate that the comprehension of the phrase: _"Person P borrows object O from donor D"_ involves at least nine implications such as _"Person D had possession of object O"_, _"D knows that P has O"_, and _"D gave permission for P to have O."_ These implications are drawn from semantic memory. Drawing implications is essentially a decomposition process in which concepts are decomposed into more basic underlying concepts. For example: "The semantic elements of GIVE are DO, CAUSE, and TRANSFER; other words such as _trade_ or _sell_ require additional semantic elements for their use, and are, therefore, more complex semantically" (Kintsch 1977, p. 351).

Implications may be compared with memory for events that are stored in episodic memory. For example: "I was with D when he lost object O" or "I saw P take object O without asking D." Comparison with any of these episodes would require additional elaboration to allow the statement to be comprehended. It may require updating the schema relating the loss of object O: "D must have found object O since I last saw him." It may instead result in the rejection of the veracity of the phrase, or its timing "I don't believe P borrowed object O at all," or "Perhaps P borrowed object O before D lost it." Another possibility is the rejection of one or more of the implications of the concept _borrow_. This would be unlikely for well-known concepts, but would be a possibility in the earlier stages of concept learning. "We propose that information received early in the growth of a cognitive category is more influential than later information in shaping that category and is less shaped _by_ it" (Jones and Gerard 1967, p. 141).

For communication to occur successfully, the speaker and listener must agree on the implications of words. Consider the kidnappers in The Princess Bride (Goldman 1973): the Sicilian (the leader and thinker), the Turk (the man of force), and the Spaniard (the deft artisan with a sword). They have kidnapped Princess Buttercup (the most beautiful woman in the world) and are sailing across Florin Channel to safety. The Sicilian, feeling cocky, speaks up:

> "We are miles ahead of anybody and safe, safe, safe."
>
> "No one could be following us yet?" the Spaniard asked.
>
> "No one," the Sicilian assured him. "It would be inconceivable."
>
> "Absolutely inconceivable?"
>
> "Absolute, totally, and in all other ways, inconceivable," the Sicilian reassured him. "Why do you ask?"
>
> "No reason," the Spaniard replied. "It's only that I just happened to look back and something's there."

And there was a ship behind them with a lone man dressed in black at the tiller. In discussing the man's purpose, the Sicilian concludes:

> "...he is definitely not, however much it may look like it, following us. It is coincidence and nothing more."
>
> "He's gaining on us," the Turk said.
>
> "That is also inconceivable," the Sicilian said. "Before I stole this boat we're in, I made many inquiries as to what was the fastest ship on all of Florin Channel and everyone agreed it was this one."
>
> "You're right," the Turk agreed, staring back. "He isn't gaining on us. He's just getting closer, that's all."
>
> "It is the angle we're looking from and nothing more," said the Sicilian.

When they have scuttled their ship at the base of the Cliffs of Insanity, and Fezzik, the Turk, has climbed a rope seven hundred of the

thousand feet up the Cliffs, carrying the other three, the man in black begins climbing after them.

> "I can feel him," Fezzik said. "His body weight on the rope."
>
> "He'll never catch up!" the Sicilian cried. "Inconceivable!"
>
> "You keep using that word!" the Spaniard snapped. "I don't think it means what you think it does." (pp. 88-92)

Effects of elaboration:

Associating an incoming message with existing schemata has several effects. One effect is to provide a means of more fully understanding the meaning and implications of the message. The existing schemata provide what Ausubel calls "ideational anchorage" (1963, p.220) "...and make possible the perception of insightful relationships (p. 218). "The meaning of a proposition ...is the set of hypothetical statements one can make about attributes or consequences related to that proposition" (Bruner 1957, p. 126). These attributes or consequences of a proposition are inferred from memory for relationships or events. "According to this view, the cognitive system interprets each new situation as being similar to some previously encountered situation, except for specifically noted differences" (Norman and Bobrow 1976, p. 130).

A second effect is to provide the organization that allows the message to be incorporated into the overall cognitive structure. Readers have a greater likelihood of processing propositions that relate to propositions processed earlier (Kintsch 1975, p. 98). Norman (1976, p. 224) notes that such a system is economical as it does not necessitate learning new material from scratch, but just requires the relating of it to previously existing schemata. "The more we comprehend the past, the better we apprehend today" (Cherry 1966, p. 73).

A third effect is the distortion of the incoming message. The constructive aspect of the elaboration process results in additions to and deletions from the initial text base. "...The cognitive system uses the first parts of the incoming information to select an appropriate, existing schema as a basis for interpreting the stimuli; as more and more information accumulates, either the new information is fit into an appropriate spot in the existing schema or the schema is modified" (Norman and Bobrow 1976, p. 130). This process corresponds to Piaget's "assimilation" and "accommodation" processes in which information is admitted to the cognitive structure and then changes are made in the message or in the existing schemata to make the two compatible (Mayer 1977, Ch. 8). Atkinson and Shiffrin (1968, p. 115) have called elaboration a coding process--"a select alteration and/or addition to the information...as a result of a search of the long-term store."

The important point is that this process is not a passive one resulting in accurate representation of incoming messages, but an active, constructive one which produces a representation that is highly dependent on prior information structures (schemas) that are stored in memory (see Howe 1970, p. 218; LaBerge and Samuels 1974, p. 320; Cofer 1976, p. 10; Posner 1973, p. 36; Norman and Bobrow 1976, p. 130; and Brockway, Chmielewski, and Cofer 1974, p. 207).

Message-Belief Discrepancy: Consistency and Change in the Conceptual Structure

A person who is confronted with a persuasive message is likely to find that portions of the message are discrepant with beliefs that he/she holds. Both of the discrepant positions cannot be correct simultaneously. The recipient could believe both, change his/her

position, or distort the meaning of the message. The choice involves two recurring themes regarding the overall conceptual framework in memory: there is a need for internal consistency and there is a need to update the system to include adaptive changes necessary for survival in a changing environment.

The Need for Consistency:

The need for internal consistency is one that has been heard from many quarters. Hovland, Janis and Kelley (1953) mentioned that as early as 1925 Lund was postulating a striving for consistency (Lund 1925). In discussing Piaget's stage-independent views of cognitive structure, Mayer said: "We have a constant need for our representation of the world to be well-organized, internally consistent, and orderly..." (1977, p. 175). "The basic principle of human information processing is that the cognitive system attempts to create a cohesive structure of the data presented to it" (Norman and Bobrow 1976, p. 130). Rokeach (1968a, p. 164) considered that "...consistency with self-esteem is probably a more compelling consideration than consistency with logic or reality." McGuire echoed the assumption of consistency: "I assumed that the conceptual system at any moment was highly interconnected and also in a state of internal harmony that might reasonably be called 'consistency'" (1968, p. 140). "...Inconsistency...threatens basic needs for a stable self concept and a coherent and predictable environment" (Kelman and Baron 1968, p. 332).

Norman (1976, p. 73) stated: "the aim of cognitive processes is to form a meaningful interpretation of the world. Sensory information at any moment must be gathered together and interpreted in terms of a coherent framework."

Internal coherence or consistency within the cognitive structure minimizes conflicting response tendencies to stimuli occurring in the environment. The result is an "<u>unequivocal behavioral orientation</u>... When the time comes to act, the great advantage of having a set of coherent internally consistent dispositions is that the individual is not forced to listen to the babble of competing inner voices" (Jones and Gerard 1967, p. 181).

These needs for consistency suggest that if a message contains information inconsistent with the existing cognitive structure, the elaboration process would result in either the modification of the incoming information or of the prior structure. The problem is <u>which</u> to modify to attain consistency.

Modifying the prior structure has a cost, as changes in one part will result in compensating changes throughout the system (Rokeach 1968b, p. 22). Because the implications of changes in central or more general concepts extend to subordinate concepts, the more central or general the concept changed, "...the more widespread the repercussions in the rest of the belief system" (Rokeach 1968a, p. 3; see also Ableson 1968, p. 716). It has been suggested that some discrepancy will be tolerated without requiring any change if the discrepancy is below a threshold level. McGuire calls this concept "spatial inertia" (McGuire 1968, p. 158; see Sherif and Hovland 1961 for the related concept, <u>lattitude of acceptance</u>).

Abelson (1959) suggested that reality, as well as the overall belief system, determine the boundaries of distortion. Distortions which result in beliefs which are contradictory to the evidence of reality as the recipient confronts it are probably difficult to maintain.

Another constraint on cognitive distortion is reality as it is socially defined. Social comparison theory (Festinger 1954a) considers an individual's dependence on others for a concensual representation of reality.

The cost of changing the prior structure is increased if the change will result in public awareness of the temporal inconsistency of these beliefs (Tedeschi, Schlenker and Bonoma 1971; see also Hovland, Janis and Kelley 1953, p. 127).

The Need to Update the Cognitive System:

Modifying the content of the message also has associated costs. With a changing environment "...we have a need to bring in new information which will disrupt the internal organization but will help us better to survive and to adapt to the reality of the external world" (Mayer 1977, p. 175). Maintaining the existing structure in the face of conflicting evidence is maladaptive and would leave the individual with inappropriate response schema.

Some studies of dissonance theory examine the problem of persons confronted with information that is dissonant (inconsistent) with the individual's existing cognitive structure. Kiesler, Collins, and Miller noted that "...one of the basic postulates of the theory is that a person will actively attempt to avoid situations and information that would increase dissonance" (1969, p. 223). Yet some studies show that the usefulness of the potentially dissonant information moderates the avoidance tendency (see Sears and Freedman 1967). This is consistant with the adaptive value of updating one's cognitive structure.

However, the negative consequences of inappropriate responses based on outmoded or unsuitable schemata are not always greater than the cost

of changing the cognitive structure (see Cohen and Goldberg 1970; and Cohen and Houston 1972). If the cost of changing the cognitive structure exceeds the perceived risk accompanying an inappropriate response, then the discrepant information should be avoided or rejected (see Kelman and Baron 1968).

The Effect of Uncertainty in Discrepancy Resolution:

Another dilema confronting a person who must resolve a message-schema discrepancy is which is more likely to be correct, the incoming information or the schema. This is equivalent to the hypothesis testing dilemma facing researchers. They can make an error by saying a relationship exists when it doesn't or they can make an error by saying that a relationship doesn't exist when it does (see Mayer 1977, Chapter 2 for a hypothesis testing approach to concept learning). Thus, a person can make an error by rejecting incoming information when it is true or by modifying his cognitive structure when the incoming information is false. The resolution of this dilemma should depend on the confidence the person has in his own representation compared to the confidence he has in the source of the message. Kelman (1961) proposed that the perceived credibility of the source depends upon the trustworthiness and the expertise that the source is perceived to have. Of course, the content of the message itself (i.e., phrasing, logic, arguments) also affects the perceived likelihood of its correctness.

Isolation of Inconsistent Beliefs within the Cognitive Structure:

Kelman (1961) suggests an interesting property of a person's cognitive structure that is in line with the context-dependent-description-view of memory schema suggested by Bobrow and Norman (1975). This property is the ability for someone to keep responses separated by the

context to which they are appropriate. Thus, inconsistencies refer only to the subsystem within which the response resides. One response may be inconsistent when viewed within the person's total value system, yet may be perfectly acceptable given that it is expressed within the context of gaining the favor of someone with the reward and punishment power. Similarly, some responses may only be consistent within the context of establishing a self-defining relationship between the individual and a group or another individual. Kelman (1961) suggests that as the generality of the affected schema increases to include the person's value system, credibility of the source of information becomes the criterion for judging the veracity of the information. Kelman states that the relevant source variables differ depending upon the context within which the schema is relevant. If the schema is relevant within the context of seeking rewards and avoiding punishments, the reward and punishment power of the source is the relevant variable. If the context is one in which the schema establishes a self-defining relationship, then the attractiveness of the target individual or group is the relevant variable. Tedeschi and Lindskold (1976, p. 336) add status as a relevant source variable.

Given these source variables, a person has another alternative when confronted with information that is inconsistent with his cognitive structure. In addition to the elaboration options of modifying the incoming information or the existing schema, he can also deprecate the power, attractiveness, credibility, or status of the source. This could justify maintaining the prior structure in the face of discrepant information. Another option would be to switch relevant contexts: "Sure, I said that; but I don't really believe it. I just said it so he

would...etc." Or, "I just behave that way around my fraternity brothers."

Message Recall Over Time: A Structural View of Memory

After a person has been exposed to a message, what will be recalled? Many researchers of memory have used a classification scheme that assumes three separate memories: iconic or short-term sensory storage; short-term or active memory; and long-term storage. (See Atkinson and Shiffrin 1968; Neisser 1967; and Lindsay and Norman 1977.) This approach assumes that information longevity depends on the memory in which it resides--only a few seconds in iconic storage; about 20 seconds in short-term memory unless the material is rehearsed; and an indefinite time in long-term storage. It is further assumed that the information has a different form in the three memories: a sensory image, corresponding to the mode of presentation (i.e., acoustic, visual, tactual, kinesthetic) in short-term sensory storage; primarily verbal in short-term or active memory; and primarily conceptual (semantic) in long-term memory. It is postulated that information can be accessed from long-term memory (LTM) and rehearsed in short-term memory (STM) along with information that has been converted from iconic to verbal form. Rehearsal is then considered a necessary activity to transfer the contents of STM to LTM. "We shall use the term memory cell to refer to a set of memories which are activated and then stored together within the long-term memory of the subject. The items within a memory cell may include those which have been presented to the subject from the external world and those associated thoughts which are activated from the subject's own long-term memory" (Posner 1973, p. 29).

Message Recall Over Time: A Processing Viewpoint

Another approach to memory phenomenon eliminates the strictures associated with presuming three separate structures. This other approach views memory persistence as a result of the extent of processing, rather than as a result of which memory structure the information resides in. As originally formulated, this approach considered memory retention to be a function of depth of processing: "...processing levels may be more usefully envisaged as a continuum of analysis. Thus, memory, too, is viewed as a continuum from the transient products of sensory analyses to highly durable products of semantic associative operations" (Craik and Lockhart 1972, p. 676). In 1975, Craik and Tulving offered a slightly different view: "A more promising notion is that retention differences should be attributed to degrees of stimulus elaboration rather than to differences in depth" (p. 279). This view has empirical support for recall of words (Craik and Tulving 1975) and for sentences (Triesman and Tuxworth 1974; Mistler-Lachman 1974). Norman and Bobrow (1976, p. 128) offer a similar view. "A description based on the physical features of the item works only as long as those physical cues are relevant to the addressing structure of the memory.... Some description of the contents or meaning of the items appears to be necessary for future retrieval." (See also Kintsch 1977, p. 327).

Slobin offers an interesting footnote (1974, p. 129) that relates the effect of language on the focus of research efforts and is relevant to the competing views of memory: "In a similar vein, consider the many nouns used by psychologists--'mind', 'behavior', 'cognition', 'rules'...

Our vocabulary can lead us astray here as well, promoting an endless search for psychological 'entities' where we ought to seek understanding of processes and dynamics, equilibrium and disequilibrium, and other more 'verb like' notions."

Whichever approach is most appropriate, the conclusion is the same: elaboration of an incoming message by relating it to existing memory schemata facilitates long-term retention, and memory for physical properties of the stimulus will be short unless those properties were unusual. (See Kolers and Ostry 1974; and Bobrow and Norman 1975, p. 144).

The Role of Elaboration in Memory:

A cue representing any part of the composite, elaborated memory unit will allow access to the entire unit (Craik and Tulving 1975, p. 291; Keele 1973, p. 46; Hayes-Roth 1977, p. 262). Unless physical features of a stimulus are unique, physical cues access so many memory units that a single memory unit cannot be accessed unambiguously. On the other hand, extensive elaboration provides many cues that may access the memory unit. Simultaneous description of the memory unit by mutiple associations also makes that unit more highly differentiated and makes designation of the unit less ambiguous when recalled. "Each 'deeper' description disambiguates the item further from a wider range of possible retrieval contexts" (Norman and Bobrow 1976, p. 129; see also Kintsch 1977, p. 361).

Elaboration also organizes information within memory. "An integrated or congruous encoding thus yields better memory performance, first because a more elaborate trace is laid down and, second, because richer encoding implies greater compatibility with the structure, rules,

and organization of semantic memory. This structure, in turn, is drawn upon to facilitate retrieval processes" (Craik and Tulving 1975, pp. 291-292; see also Keele 1973, pp. 44-45).

The initial processing of a message thus determines whether the content can be unambiguously recalled and the contexts within which the content may be recalled (Cofer 1977, p. 336; LaBerge and Samuels 1974, p. 320; Craik and Tulving 1975; Craik and Lockhart 1972, p. 678; Lindsay and Norman 1977, pp. 374-375; Bobrow and Norman 1975, p. 133; Shiffrin and Schneider 1977, p. 157; Posner 1973, p. 30; Norman 1976, p. 127).

What people recall is not what occurred--"...not what was 'out there' but what they did during encoding" (Craik and Tulving 1975, p. 292). "...The record is not of the order of external happenings, but rather of the order of our own thoughts at the time" (Posner 1973, p. 29). This event-based memory can also influence semantic memory, resulting in what Posner calls "parallel storage." "New input is always embedded within the context of contiguous external and internal events which are active in memory....When time is available, considerable cross-filing and referencing of stimulus information may take place. Thus new input can be incorporated in the hierarchical structures..." (Posner 1973, p. 36).

However, as similar events occur in more and more contexts they can take on more of a conceptual and less of an event-based representation. "Although the stability of meaningful material is initially enhanced by anchorage to relevant conceptual foci in the learner's cognitive structure, such material is gradually subjected to the erosive influence of the conceptualizing trend in cognitive organization. Because it is more economical and less burdensome to retain a single inclusive concept

than to remember a large number of more specific items" (Ausubel 1962, p. 218).

Reconstructive Processes in Recall:

Earlier in this chapter, the constructive nature of the encoding process was discussed. However, recall is also reconstructive. That is, if individuals are called upon to recall a situation or story, they recall the overall theme or gist and fill in the detail with "plausible" reconstructions (Kintsch 1977, p. 363). Bartlett (1932) wrote a very influential book in which he reported numerous studies of remembering along with his "Theory of Remembering" (Chapter 10). This reconstructive phenomenon was very prevalent in his observations, and "In the many thousands of cases...collected...literal recall was very rare" (Bartlett 1932, p. 204) (See both Mayer 1977, p. 112; and Norman 1976, p. 223 for a discussion of Bartlett's work.) Bartlett does not see this as maladaptive: "In a world of constantly changing environment, literal recall is extraordinarily unimportant" (Bartlett 1932, p. 204).

Recall for prose is not limited to reconstruction. It is "...neither reproductive or constructive nor reconstructive, but all three..." (Kintsch 1977, p. 363). Reconstruction, as does construction, appears to be largely an inferential process. Franks and Bransford (1972) found that sentences that could be inferred from a presented text were recognized as having appeared even when they had not. Griggs (1976) reports similar findings. (See also Dooling and Cristiaansen 1977, p. 428.) Kintsch (1977) reports that reconstruction increases as the time since the presentation of a message increases. Reproduction decreases during this period.

Howe (1970) reported another interesting time effect with repeated exposures. Errors that appeared in initial attempts to reproduce a prose passage were not corrected when Ss were given playback of the prose passage immediately after the reproductive attempt. Their errors persisted in subsequent reproductive attempts. "In the case of verbatim recall scores, even incorrect additions are, on the average, two or three times more likely to recur on the succeeding trial than are previously nonrecalled items to be reproduced correctly" (Howe 1970, p. 218).

Reconstruction and generalized schemata:

Reconstruction may often depend on frameworks that exist prior to exposure to a message. "The most dramatic instances of reconstructions that we have observed in our work occurred when we gave subjects texts to read about topics [with] which they were quite familiar..." (Kintsch 1975, p. 99). In addition to frameworks based on personal familiarity, many frameworks are culture-based. (See the earlier discussion of the influence of these frameworks in the comprehension process.) Culture-based frameworks exist for many types of communications in a society. They consist of a set of expectations of the order and content of the communication. In his discussion of natural narratives, Kintsch (1975, p. 106) calls "...these expectations a narration schema." Thus, there could be lecture schemata, conversation schemata, scientific article schemata, story schemata, advertisement schemata, etc. Given the framework and the topic, Ss could probably do an admirable job of reconstructing the message with no exposure.

As it is, reconstruction probably contaminates many attempted measures of verbatim recall. "If a person remembers the concepts

communicated in a text and how they were related as well as some stylistic features of the passage...it would not be difficult for him or her to produce in recall a number of sentences which look as if they have been taken verbatim from the passage but which in fact have been constructed from knowledge of the gist of the passage in the same way as those recalled sentences whose structures deviate from the style and form of the input" (Cofer, Chmielewski, and Brockway 1976, p. 197). The greater the familiarity with the material presented and the more standardized the cultural framework for the type of passage, the more likely is reconstruction to occur. Increasing the delay since presentation and the length of the text should also increase reconstruction (Kintsch 1977, p. 364). It seems that this effect could be measured by asking control Ss to produce what they think a message would have contained, given a topic and source. The content of their responses could be examined to determine the underlying propositional base as suggested in the earlier discussion of comprehension (Kintsch 1974 and Kintsch 1977). This propositional base could be compared with that extracted from protocols collected from Ss exposed to the message. Both of these propositional bases can be compared to the text base of the message.

In summary, recall can be expected to be strongly influenced by elaboration and laced with reconstructions to fill in any bare spots in the propositional framework that the person holds for the message. Since elaboration is so crucial to the communication process, anything that interferes with elaboration will have a drastic effect upon retention of a message. Because of this importance, interference and processing limitiations will be discussed next.

Processing Limitations

Many persuasion studies which have presumed the intervening role of subvocalizations in the communication process have utilized distracting tasks to partially preoccupy the message recipient and, thereby, prevent subvocalizations which may be counter to the intent of the message. To properly evaluate these studies, the limitations on message processing must be examined.

Attentional Requirements for Message Processing:

Examination of the problem of interference requires an understanding of different types of processing--those that require attention and those that do not. Many schemata in memory contain entire procedures that, when activated, can run their course as a standardized response. Other responses are the result of the activation of new, unique combinations of schemata. Standardized responses are lower level responses such as word recognition or the extraction of the physical features of stimuli. These lower level responses are automatic. That is, they do not require attention (LaBerge and Samuels 1974, p. 295; Craik and Lockhart 1972, p. 672; Shiffrin and Schneider 1977, p. 159). Their activation is callled "effortless" (Posner 1973, p. 40). Higher-level, non-standardized responses require attention--they are "effortful" (Posner 1973, p. 40).

Attention has been used to refer to several different concepts in the cognitive psychological literature: alertness; selectivity of information source; and a "limited central processing capacity" (Posner and Boies 1971, p. 391). It is the processing capacity limitation concept which applies here. "Basically, the literature from

experimental psychology on attention indicates that the central high-level cognitive mechanisms have a limited processing capacity" (Bobrow and Norman 1975, p. 138). "Task performances that deteriorate under simultaneous conditions are said to demand attention" (Keele 1973, p. 4). So attention is defined in terms of interference. If two processes interfere with each other, they both require attention or a part of the limited processing capacity.

Automaticity, in turn, is defined in terms of attention. "Our criterion for deciding when a skill or subskill is automatic is that it can complete its processing while attention is directed elsewhere" (LaBerge and Samuels 1974, p. 295). Automaticity increases with the familiarity of the stimulus and the number of repetitions of the response (Shiffrin and Schneider 1977, p. 159; LaBerge and Samuels 1974, p. 320).

Degree of elaboration and attention:

The greater the depth of processing, the less likely is the processing to be automatic and, therefore, the more likely it is to require attention (Bobrow and Norman 1975, p. 143). Thus, while accessing the meaning of familiar words in memory may be automatic, the deeper elaborative processing necessary to extract the propositional underpinning of a sentence requires attention (Posner 1973, p. 139).

How extensive is the interference between two processes that require attention? "In general, one does not err much by assuming that only a single high-level task can be performed at any given time" (Bobrow and Norman 1975, p. 139). A person cannot carry out the elaborative process and another high-level process at the same time (Posner and Boies 1971, p. 391).

Speaking interferes with listening unless the phrases spoken have little content as in "carrier phrases such as 'Hello' and 'This is'" (Broadbent 1952, p. 272). However, it is possible to perform one high-level process and others which are automatic at the same time (LaBerge and Samuels 1974, p. 295). Pintner (1913) used vocalized counting to occupy the speech musculature and, thereby, prevent articulation of the text during reading. "At first there were many hesitations and interuptions, but gradually, the process became automatic" (Pintner 1913, p. 140). Thus, his Ss were able to read "silently" while speaking the numbers. (See also Sokolov 1969, p. 539.)

Interference and elaboration:

Shadowing an oral message, on the other hand, requires a great deal of attention. While listening to a tape recorded reading of a prose passage, "...the listener is instructed to repeat what he hears concurrently, in a subdued or whispered voice" (Cherry 1966, p. 281). So much attention is required to perform this task that the content of simultaneously presented messages cannot be recalled. In fact, the Ss may not be able to state if the the non-shadowed material was in English or in some other language (Posner 1973, p. 126).

Memory is not just a problem for non-shadowed material following the shadowing task. The task is so demanding that Ss are unable to elaborate on the shadowed material. As a result, recall of the shadowed material is poor, "...especially if it is at all 'deep' or difficult" (Cherry 1966, p. 281; see Norman 1976, p. 19 for a related discussion).

Norman and Bobrow (1975) distinguish between data-limited and resource-limited processes in their discussion of processing capacity limitations. Processes such as signal-detection in which a person must

only decide if a light is on or not are considered data-limited and require very little of the available processing resources. However, "...the more conceptually based the decision process the more processing effort required: here, the resource limitation can severely limit the performance capability" (Bobrow and Norman 1975, p. 143).

Figure 4 shows the relationship (proposed by Norman and Bobrow 1975) between the portion of the limited processing resource devoted to a task and performance on that task. In this figure, L represents the upper limit to the processing resource. This upper limit may increase with arousal (Bobrow and Norman 1975, p. 140). The resource-performance curves for three tasks are represented in the figure--shadowing, word recognition, and tone recognition. These curves are depicted as becoming less resource limited and more data limited as they progress from shadowing to tone recognition. If shadowing were perceived by an individual as the primary task and $R_{PRIMARY}$ were allocated to it, then $R_{SECONDARY}$ could potentially be allocated to a secondary task such as word recognition or tone recognition. The figure shows the level at which these two secondary tasks could be performed if $R_{SECONDARY}$ were allocated to their performance.

Thus, performance of a conceptually-based task may leave enough available processing capacity to simultaneously perform a simple signal detection task, but not enough to carry out another high-level, conceptually-based task. (See Reitman 1974, p. 375 for the effect of "surreptitious rehearsal" on signal detection performance.) In addition to requiring a part of the limited processing capacity (space); elaborative processes also require time (Posner 1973, p. 93; Keele 1973, p. 3; Norman 1976, p. 100).

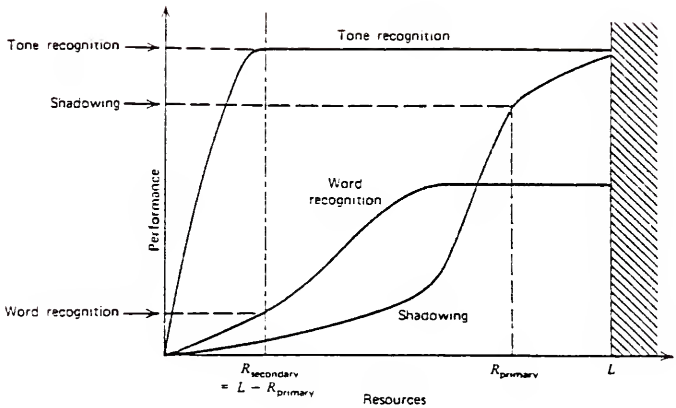


Figure 4. "Analysis of shadowing with two secondary tasks. (after Norman and Bobrow, 1975)" (from Kintsch 1977, p. 138).

The processing of a message may interfere with reception of subsequent portions of that same message. "The performance of transformations upon current input may preempt the processing capacity so that the storing of raw data is neglected" (Posner 1973, p. 146; see also Norman 1976, p. 126; and Keele 1973, p. 37). The higher the rate of presentation, the greater the deficit (Posner 1973, p. 34).

Message compression (see MacLachlan and La Barbera 1978), in which the rate of speech in a message is artificially increased, should decrease elaboration because of time pressures. Craik and Lockhart (1972, p. 670) suggest that in such circumstances, the decrease in the degree of elaboration will decrease long-term recall, but will have little effect on immediate recall. Presenting a speech-compressed message with pauses between phrases may allow intermittent elaboration and, hence, greater recall with no increase in overall message presentation time.

Interference is somewhat mode-specific. That is, two verbal tasks interfere with each other more than a spacial and a verbal task interfere with each other (Posner 1973, p. 141; Brooks 1968, p. 354).

Processing strategies for coping with limited processing capacity:

Given a limited processing capacity, how can people extract and process significant information from the melange of stimuli impinging on their senses? Several strategies are used.

One of the most intriguing strategies is characterized as the "cocktail party problem" (Cherry 1966, p. 279). This is the ability of people to select one from among many simultaneous converstaions at will. People use such cues as directionality, voice quality, context, and transitional probabilities to help them in following one of the

conversations (Cherry 1966, p. 280). However, stimuli that have high importance to an individual may intrude upon, i.e., interfere with, the conversation that is being monitored by the individual (Broadbent 1952, p. 271). That is because all stimuli impinging upon the senses are processed to determine their pertinence to the individual. Deutsch and Deutsch (1963, p. 85) suggest that only those stimuli which indicate an importance which is above the current threshold for the individual will interfere with other attention-demanding processes. This threshold may become lowered, thus exposing more stimuli, when the individual is aroused, making it difficult to process information from any of them. Deutsch and Deutsch further propose that important messages increase the general arousal level of the individual. As a result, "...messages which would not have been heeded before will command attention if they follow in the wake of a more important message" (1963, p. 85).

In addition to selectivity, switching or alternating between stimuli is another strategy that can be employed to overcome processing limitation. Since processing resources can be allocated to different stimuli by choice, an individual may switch back and forth, sampling information from alternate stimuli. How much information is lost depends upon the presentation rate and the predictability of the messages (See Bobrow and Norman 1975, p. 139; and Shiffrin and Schneider 1977, p. 156).

Conceptual processing without attention:

The lack of recall of messages presented while processing resources are directed elsewhere is somewhat misleading. There is evidence that the non-attended message modifies the perceived content of a shadowed message even when the non-attended message cannot be recalled (Makay 1973).

There is also evidence that some higher-level cognitive processes can proceed without conscious attention being directed toward them. An example is the incubation phenomenon in which problem solutions appear to evolve during a period in which attention is directed elsewhere (Slobin 1974, p. 100; see also Norman and Bobrow 1976 for an anecdotal illustration of the process). However, it may be possible that the incubation process, though unconscious, requires part of the limited processing capacity but is serially alternated with other conscious-attention demanding procedures. Since incubation is an unconscious process, it is very difficult to study, although some inconclusive attempts have been made (Fulgosi and Guilford 1968; Murray and Denny 1969; and Silveira 1971). This phenomenon and others such as the tip-of-the-tongue phenomenon led Norman and Bobrow to state: "Thus, we believe that just as the lowest levels of sensory and perceptual processing operate subconsciously by autonomous schemata, so too do the higher-level schemata form and process without conscious directions" (1976, p. 129).

Spelke, Hirst, and Neisser (1976) reported a study in which Ss were trained over seventeen weeks to extract meaning from dictation with no decrement in simultaneous reading. Their results suggest that processing capacity limitations may not be fixed, but may be expanded through specific training.

Summary of effects of interfering with elaboration:

Summing up, it seems clear that the elaboration process is one which requires part of the limited processing capacity. Except under special circumstances (see Spelke, Hirst, and Neisser 1976), only one higher-level process such as elaboration can be performed at one time.

If attention is directed toward another high-level process during message presentation, the elaboration process will be disrupted. Without elaboration, recall of propositions presented in the message will be poor or impossible (Keele 1973, p. 56). "...The concept of capacity is to be understood in terms of a limitation on processing; limitations of storage are held to be a direct consequence of this more fundamental limitation" (Craik and Lockhart 1972, p. 674). However, the selective ability exhibited in the "cocktail party problem" allows individuals to direct their attention to a message, even though confronted with another potentially distracting stimulus.

As extraneous tasks increase in complexity from simple signal detection to a high degree of elaboration, the greater will be the attentional deficit for the primary task. Non-attended stimuli may modify the individual's cognitive structure--even though presentation is not recalled.

Views of the Elaborative Process

Given that the elaboration process is so crucial to message processing, it would be very useful to monitor the elaboration process in some way. Monitoring this process would allow experimentation that could examine the effect of various independent variables on the elaboration process and the role of elaboration on persuasion. However, monitoring the elaboration process requires an understanding of its nature. Is elaboration a verbal process? Are there operational measures that can provide access to the process?

There is a long-standing disagreement about the nature of thinking. In Theaetetus, Plato equates thinking with speaking to oneself: "'The

soul when thinking appears to me to be just talking--asking questions of itself and answering them, affirming and denying...I mean, to oneself and in silence, not aloud or to another'" (Sokolov 1972, p. 34).

Watson states that: "The behaviorist advances the view that what the psychologists have hitherto called thought is in short nothing but talking to ourselves" (1966, p. 10). Definitions of rehearsal share this notion of verbal inner speech. "Define the term 'rehersal' to be conscious purposeful subvocal repetition of the items to be retained" (Reitman 1971, p. 186). "Rehearsal is a type of inner speech by which we maintain a limited amount of material in memory indefinitely" (Norman 1976, p. 100).

There are two issues that are involved here. The first is that the currency of thought is language. The second is the subjective experience of inner speech. "We must be careful to remember the distinction between language and speech" (Slobin 1974, p. 99). The relationship between inner speech and thought will be discussed in the next section.

Is Thinking a Language Process?:

Must thinking utilize words? "To divorce thought from language obviously verged on the absurd and this philosophy was completely rejected by Kant, Hegel and the linguists who followed them, Humbolt and Steinthal, for whom language was the principal actor or 'spirit', no thought being able to exist without language" (Sokolov 1972, pp. 18-19).

Others, however, see a more flexible role for language in the thought process. "We cannot necessarily put all our mental experience into [existing] words" (Cherry 1966, p. 79). "Has not every one of us struggled for words although the connection between 'things' was already clear" (Einstein 1954). "Surely we learn nonverbal experiences without any recourse to words" (Norman 1976, p. 98).

For some writers, words play the role of fixing ephemeral thoughts that may or may not have occurred in verbal form. Sokolov (1972, p. 19) relates the Hamiltonian analogy that compares thinking to digging a tunnel in sand: "Hamilton points out that when digging a tunnel in sand one cannot advance without securing every inch of the tunnel with an arch of brick, and that the same occurs in thinking: each step of the thought process has be to fixated with words, so that no forward movement of thought is possible without words." "The only way to pin down a thought before it can slip away and fly out the window is to jump on it with both verbal feet, to pin it down with language, by diagrams, or by mathematical symbols--though such language may be inadequate" (Cherry 1966, p. 79).

John Locke (1690) also relates the inequality between word and thought: "'When we begin to fix by means of words...abstract ideas... there is a danger of error. Words should not be treated as adequate pictures of things; they are merely arbitrary signs for certain ideas--chosen by historical accident and liable to change'" (Cherry 1966, p. 70; see also Cicourel 1974, p. 67 and Pylyshyn 1973, p. 7). However, the reader should recall the earlier discussion of linguistic relativity and determinism as it relates the effect of language on the codification of events (see Slobin 1974).

Cognitive representations of experience and knowledge occur as motor programs, imagery, and as language (Slobin 1974, p. 110); see also Lindsay and Norman 1977). It is likely that we respond to messages by activiating each of these on occasion. "Like all word concepts, turbulence is defined in a person's lexicon, both by stating its relationship with other word concepts and by means of appropriate

sensory imagery and motor programs" (Kintsch 1975, p. 91; see also 1977, p. 328).

Pylyshyn (1973) discusses the possibility of another cognitive "language" which may serve as a transitional representation between imagery and verbalization: "the need to postulate a more abstract representation--one which resembles neither pictures nor words and is not accessible to subjective experience--is unavoidable" (p. 5).

In summary, it appears that thinking, and thus elaboration, is sometimes a verbal process, but that is not necessarily verbal.

Can the Elaboration Process be Monitored?:

Centralists vs peripheralists:

Is it possible for a researcher to examine the elaboration process? There are two distinct viewpoints reflected in answers to this question. One is that held by the centralists and the other is held by the peripheralists. "Crudely put, the centralist theory says that thinking occurs only in the brain, while the peripheralist theory holds that thinking necessarily involves muscular movements" (McGuigan 1966, pp. 4-5). The implications are that if the centralists are correct, some thinking may take place with no measurable physical responses occurring, whereas if the peripheralists are right, thinking can be monitored by measuring the appropriate motor discharges (McGuigan 1966, p. 294). "...we shall consider the *major* (priority) definitional indicator [of thinking] as *the composite of bodily activities that necessarily occur following the presentation of a language stimulus*" (McGuigan 1973, p. 346).

Watson is the epitomy of a peripheralist (behaviorist): "A whole man thinks with his whole body in each and every part" (1920, p. 88).

Watson was exasperated with the introspectionists whom he thought were concerned with merely attacking each other's definitions and training when they should have been concerned with the introspectionist approach itself (1913). "The time seems to have come when psychology must discard all reference to consciousness; when it need no longer delude itself into thinking that it is making mental states the objective of observation" (Watson 1913, p. 163). Watson's approach was not new. Pintner (1913, p. 132) reported Stricker's (1880) view that: "Ideas of words consist of nothing else than the consciousness of the excitation of those motor nerves that are connected with the articulatory muscles."

Some of the peripheralists believe that muscular responses play an organizing role during thinking. "It may be that the speech muscle activity generates a linguistic code that is carried to the brain... [and] these afferent neural impulses may help integrate the various verbal regions of the brain" (McGuigan 1973, p. 372). Again Pintner showed that this was not a new idea as he quoted Bawden (1900): "There must be this return wave of kineasthetic imagery to select and organize the visual and auditory perceptions before they can have meaning" (Pintner 1913, p. 134; see also McGuigan 1970, p. 309). (See McGuigan 1970, pp. 321-322 for descriptions of intermediate, as well as extremist, positions in the centralist--peripheralist controversy).

Approaches to monitoring inner speech:

Historically, there have been numerous ingenious techniques invented to measure responses accompanying and indicating thinking. Most, especially the earlier approaches, assumed that thinking is a form of covert (inner) speech. This is a peripheralist viewpoint and presumes that the use of language as units of thought will be

accompanied by <u>inner speech as a locus</u> of thought (see Slobin 1974, p. 99). "From the psychophysiological point of view, internal speech is defined as latent and reduced articulation of words accompanied by heighted tonus of the speech musculature or by disconnected bursts of speech motor activity and sometimes by entirely obvious movements of the speech organs" (Sokolov 1969, p. 568). Others, such as Watson, believed that the loci of thought process extended to responses involving the whole body. "...I believe that we do [think with our whole bodies], when one considers the massive amount of evidence that implicates the eyes, arms, GSR, heart, etc., in thought processes" (McGuigan 1973, p. 374).

Measurement approaches to thought-related responses have included: a) observation, in which lip movements and audible whispering were recorded (see Pintner 1913); b) various pneumatic devices in which rubber bags or bulbs were placed in the mouth or against the larynx and resulting changes in air volume were transferred to mechanical recording devices; c) a modified glass into which the tongue was inserted--resultant movements were then mechanically recorded; d) a suction cup which was on the end of a lever was attached to the tongue and the lever amplified tongue movements (see Thorsen 1925, pp. 3-4; Lyon and Waengler 1976, p. 15; and Sokolov 1972, Ch. 2 for a description of these mechanisms).

In the 1920's Jacobson, in collaboration with two engineers at Bell Telephone Laboratories, developed a neurovoltmeter which could measure fractions of a microvolt. This was then used to monitor muscular stimulation that accompanied covert mental events (Jacobson 1973). It was found that the particular muscles stimulated were highly task-specific "...and occur in the part of the body that one would use should the response have been overtly made" (McGuigan 1970, pp. 314-315).

Many studies have been done which have utilized this type of equipment with needle, suction cup, or wire grid electrodes (see Lyon and Waengler 1976; and Hardyck, Petrinovich, and Ellsworth 1966). The specificity of response allows researchers studying covert speech phenomena to monitor speech-related musculature (see McGuigan 1973, 1966; and Sokolov 1972 for reports of numerous studies).

Other measures of thought-related events:

Additional indicator measures include electroencephalography for monitoring central nervous system events (McGuigan 1966; Krugman 1971); and pupilometry (Hess 1968), electrocardiography and electrodermal measures for autonomic events (McGuigan 1966 and 1973).

Evidence from electromyographic research:

What insight do electrophysiological measures provide about mental events? Electromyographic (EMG) indications of covert verbal activity have been found to accompany both known covert speech, listening to an oral presentations and reading (see Sokolov 1972; McGuigan, Keller, and Stanton 1964; McGuigan 1970, 1966, 1973; Lyon and Waengler 1976). Sokolov's account of when speech motor impulses occur sounds much like a description of elaboration: "They...occur when the subject analyzes the meaning of perceived speech (e.g., in picking out the main point, in reformulating the sentence, etc.)" (1969, p. 555). Here, at last, it would seem, is a measuring technique that will provide a means of monitoring the elaboration process. Two assumptions are needed: that the elaboration process utilizes covert speech and, thereby, language; and that all speech-oriented thoughts will have associated electrical activity that can be detected by the use of electromyographic equipment. These assumptions are not justified.

First, some types of elaboration may consist of non-language responses. Secondly, EMG indications do not always accompany language-based phenomena--thinking or listening. In fact, Hardyck, Petrinovich and Ellsworth (1966) were able to train Ss to extinguish subvocalization during reading by giving feedback of covert speech activity detected by "...mesh electrodes placed over the thyroid cartilage" (p. 1467). This negative result was also found with observational and mechanical measurement approaches (see Slobin 1974, p. 118; and Thorsen 1925).

Determinants of EMG indications of thought processes. The registration of covert verbal activity seems to depend upon how difficult or novel the thoughts are and whether the response is one that has become automatized (Sokolov 1969, p. 569). For listening, a "...significant difference was found for covert lip activity on graded verbal stimuli. As test passages increased in difficulty, mean passages increased in difficulty, mean phasic muscle activity for the lip showed a corresponding increase" (Lyon and Waengler 1976, p. 7).

In reading, textual difficulty also affects measured covert speech activity (McGuigan 1973, p. 364; 1970, p. 317). It also appears that there are individual differences in exhibited covert speech activity during language tasks. More proficient readers exhibit less covert speech activity (McGuigan 1973, p. 364). "...One's ability to use language may be an important factor in determining whether covert speech activity at the articulators is present for extracting meaning" (Lyon and Waengler 1976, p. 20). LaBerge and Samuels' (1974, p. 306) discussion of reading suggests why this may be the case: "If he is reading easy material at a fast pace, he may select as visual units words or even word groups; if he is reading difficult material at a slow pace, he

may select spelling patterns and unitize these into word units at the phonological level." This may be a case in which the musculature or acoustic imagery can aid in the thinking process, just as when reading very difficult material, it may help to read the material aloud.

Autonomic indicators of thought processes:

An autonomic indicator with thought monitoring potential is pupil dilation (Hess 1968). Pupil dilation has been found to accompany emotional responses as well as the solving of difficult problems (see Blackwell, Hensel, and Sternthal 1970 for a review). However, pupil dilation seems to be too generalized to be a useful indicator. "The authors' current impression is that pupillary dilations probably accompany any substantial increase in mental activity, regardless of the pleasantness of this activity" (Blackwell, Hensel, and Sternthal 1970, p. 18).

Another measure that may be useful in monitoring central nervous system events accompanying thought processes is the electroencephalogram (EEG) (McGuigan 1966, p. 294; Krugman 1971).

In summary, it appears that thought processes sometimes utilize language and sometimes result in measureable concomitant psychophysiological responses.

Verbal report of thought processes:

One remaining avenue of monitoring the elaboration process is to use the Ss as observers of their own thought processes. Through introspection, the Ss try to recall the thought process and to give a verbal report of it. This record is called a verbal protocol (Lindsay and Norman 1972, p. 502). For this to be an accurate rendering of what occurred, the process must: a) be observable by the subject; b) be

recognized as a significant part of the thought process; c) be verbalizable; d) be recallable.

Self-awareness of one's thought processes:

For a mental event to be observable by the Ss, it must be one of which he is aware. There is some evidence that some thought processes are not observable. One recurring phenomenon is incubation. There are many introspective reports of problem solutions that have "popped into a person's head" after a period in which no conscious attention was directed to solving the problem (see Norman and Bobrow 1976, pp. 115-116 for a personal account of the phenomenon). In these cases, the process was not observed by the Ss and could not be described.

Another example is "highway hypnosis" (Williams 1963). In this situation, Ss report driving for miles without being aware of the myriad decisions that must have been made to keep the car under control. These are both examples of automaticity. The first is a high-level, conceptually-driven process and the second is a low-level, data-driven process. Norman and Bobrow (1976, p. 129) suggest that the first (incubation) is a result of high-level schemata communicating directly with each other within our cognitive structure, without conscious attention. The second, a series of automatized responses, also occurs without conscious attention. However, lack of awareness means that no report of the processes can be made--that there is no "...introspective knowledge of them" (Natsoulas 1970, p. 92; see also Broadbent 1958, p. 52).

Recognition of significant mental events. Since some processes are not observable, those that are may take on undeserved significance. "While most psychologists are willing to concede that not all important

psychological processes and structures are available to conscious inspection, it is not generally recognized that the converse may also hold: that what is available to conscious inspection may not be what plays the important causal role in psychological processes" (Pylyshyn 1973, pp. 2-3).

In the 1800's introspectionist researchers trained Ss to recognize significant mental events and to describe their thought processes. However, the results of the introspective process depended upon the particular training that the Ss had received. In attacking the technique, Watson (1913, p. 163) sums up the typical diatribe between introspectionists: "If you fail to reproduce my findings, it is not due to some fault in your apparatus or in the control of your stimulus, but it is due to the fact that your introspection in untrained." Nevertheless, the fact remains that Ss may not report some mental events that played a role in the elaboration process. Training, on the other hand, may only standardize reports (see the discussion of linguistic relativity and determinism) and lead to reports of events that the Ss think <u>should</u> have occurred. "...The instructions can impose on the reports an artificial, non-spontaneous character including support for the preconceptions implicit in the instructions...[the] instructions may [also] interact with individual characteristics (like intelligence)" (Natsoulas 1967, p. 253).

<u>Verbalization of mental events</u>. Another requirement for accurate reporting of mental events is that they be verbalizable. Slobin (1974, p. 100) discusses the tip-of-the-tongue phenomenon. In this situation an individual cannot verbalize a word to express a concept. Yet it is apparent that the thought of the word has occurred since Ss recognize

the word if it is presented and can often tell what letter it begins with and how many syllables it has!

People often have difficulty expressing thoughts in words when the thoughts did not occur in words. Images, emotions, and intuition must be transformed before they can be reported (see Slobin 1974, p. 101; Cherry 1966, pp. 71, 78-79; and Cicourel 1974, p. 67).

Recall of mental events. Another limitation to the accuracy of verbal reports of mental events is the inability to recall exactly what transpired. This limitation was recognized as early as 1882 by John Stuart Mill. Boring (1953, p. 171) reports: "It is not strictly immediate, Mill thought, for it involves memory--immediate memory, perhaps; yet immediate memory is not the datum itself and comes with a chance for error in it" (see also Bakan 1954, p. 109).

The very process of reporting what had occurred would prevent maintenance rehearsal of other thoughts that had occurred. The same interference could result from listening to the instructions for reporting the episode. Thoughts would, thus, have to be retrieved from memory or reconstructed (see the earlier discussion of reconstructive process in memory). The individual could work backwards from the present belief structure to infer and reconstruct the elaborative responses that must have occurred to reach that state (see Bartlett 1932, p. 187).

Boring describes the typical introspective episode in Titchener's labs in the 1800's: "It could take twenty minutes to describe the conscious content of a second and a half and at the end of that period the observer was cudgeling his brain to recall what had actually happened more than a thousand seconds ago, relying, of course, on inference" (1953, p. 174).

The greater the elapsed time between the thought episode and the report, the less likely is verbatim recall of any verbal portion of that episode, and the greater is the likelihood of a rephrasing of the underlying propositional base of that thought sequence (Kintsch 1977, p. 364).

The major shortcomings of introspective accounts are that the report is distinct from the event (Bakan 1954, p. 107; Natsoulas 1967, p. 251; Boring 1953, p. 184), and that the Ss as an observer is a source of error: because the event was not observable; because the Ss did not recognize significant events; because the Ss cannot verbalize his thoughts; or because he cannot accurately recall the character of his thoughts.

Conclusions regarding verbal reports of thought processes. Is it then necessary to agree with Nisbett and Wilson? "The evidence reviewed is then consistent with the most pessimistic view concerning people's ability to report accurately about their cognitive processes....Such reports, as well as predictions, may have little value except for whatever utility they may have in the study of verbal explanations per se" (1977, p. 247).

It is unlikely that any verbal protocol will exactly mirror the elaboration process. Its value lies in its ability to provide insight into the product of that elaboration. Verbal protocols that reflect reconstructions based on inferences from the present state or upon rephrasings of the underlying propositional base of elaborative responses can provide evidence of the nature in which a message was coded--of the underlying propositional representation of the content of the message as it was perceived by the receiver.

CHAPTER 3

THE INTERVENING ROLE OF SUBVOCAL RESPONSES IN THE PROCESSING OF PERSUASIVE COMMUNICATIONS

There is one overriding difference between the studies of message processing performed within the realm of cognitive psychology and those under the aegis of social psychology and marketing. Cognitive psychological studies have been concerned primarily with delineating processing activities and the limitations on these activities, while social psychology and marketing have been primarily concerned with how these factors affect persuasion. This chapter will present an elaboration model of message processing based on the processing principles which were presented in the second chapter. Then studies of persuasion will be positioned vis-a-vis the relationships portrayed in the model. The operationalization of predictor and criterion variables utilized in the various studies will be examined for appropriateness and validity. Following this, the evidence from these studies will be examined for their consistency with the elaboration model.

The conceptualization of elaboration draws on a definition proposed by Tulving and Madigan (1970): "Elaboration coding refers to the storage of additional nonredundant information with the verbal unit to be remembered" (p. 462). They used this definition to distinguish elaboration coding from the use of various mnemonic systems which contain information redundant with the unit being coded. This second type of coding they called "substitution coding" (p. 462). Tulving and Madigan considered that "coding operations" could be viewed as

synonymous with "differential covert responses that subjects are assumed to make to verbal items, implicit associative responses [and] mediating responses..." (p. 461). (The reader will recall the distinction between maintenance rehearsal and elaborative rehearsal from the discussion of the elaborative process in Chapter 2.)

Craik and Lockhart (1972) used this concept in their discussion of levels of processing. "After the stimulus has been recognized, it may undergo further processing by enrichment or elaboration. For example, after a word is recognized, it may trigger associations, images or stories on the basis of the subject's past experience with the word" (p. 675).

It is in the same sense that message elaboration is used in the present discussion. As such, it does not apply to the processing of a message to the "surface structure" (see Chapter 2) level, but rather additional associative responses which link the message content with other schemata within the message recipient's cognitive structure. These elaborative (associative) responses could be made when the message is first processed in an externally generated fashion or at a later time in an internally generated fashion.

Persuasion will be considered as the shifting of a person's cognitive structure from the positions held prior to the communicative episode to positions closer to those intended by the communicator as a result of the communicative episode. The person's cognitive structure is considered to be all of the schemata held in memory: values, beliefs, goals hierarchy, episodes, semantic knowledge and motor programs. Persuasion, thus, occurs when there is a change in any of these areas, in the direction intended by the communicator, resulting from a communicative incident.

An Elaboration Model of Message Processing

A message processing model which allows the examination of persuasion studies can be developed from the theoretical background examined in the previous chapter. The model incorporates the following propositions:

1. Processsing of messages involves a progression from physical feature extraction to the development of a fully elaborated propositional structure. However, elaboration need not wait until all lower level processing has been completed (Norman and Bobrow 1975, p. 45).

2. Processing may stop at any level.

3. The length of time the mental representation of the message will be retained depends on the extent to which it was processed--more extensively processed messages will be retained longer.

4. The greater the discrepancy between the implications of the message and the cognitive structure of the recipient, the greater should be the motivation for elaboration. However, the urgency for elaboration is probably moderated by other variables, such as communicator credibility, centrality of the issue, etc. The <u>extent</u> of elaboration probably depends upon the ease with which the discrepancy can be resolved.

5. Processing requires both time and space (as a portion of the limited central processing capacity).

6. Activities which compete for the time or space necessary for processing will reduce the extent to which a message is processed.

7. Reducing opportunity for elaboration should affect long-term

recall more than short-term recall.

8. When processing a message, the content and operations of elaboration will be determined by <u>message structure</u> (expectations elicited early in the message will affect processing of later portions in a top-down fashion); by the <u>source</u>; by <u>past experience</u> of the recipient (familiarity with the message); by the <u>cognitive structure</u> to which it is related; and the restrictions on the portion of the <u>limited processing capacity</u> that can be allocated to the processing.

9. How closely the post-message cognitive structure (the recipient's text base) approximates that intended by the speaker or writer depends upon: a) the degree to which the implications of the words and phrases used in the text are shared by the communicator and recipient; b) the content of elaborations made by the recipient and; c) the initial position of the recipient relative to that of the source. A very simplified elaboration model of message processing that incorporates these characteristics was presented in Chapter 1 and is reproduced here for the reader's convenience (Figure 5).

The elaboration process occurs within a context and may be affected by various contextual variables: message source; situation; surrounding environment--such as background noise; and concurrent tasks performed by the individual. The <u>nature</u> of elaboration and the <u>motivation</u> to elaborate is also influenced by the <u>prior cognitive structure</u> extant at the time of message onset. Another influence on message elaboration is the message itself, that is, <u>message variables</u> such as argument order, grammar, speed of presentation, signal quality, or the inherent

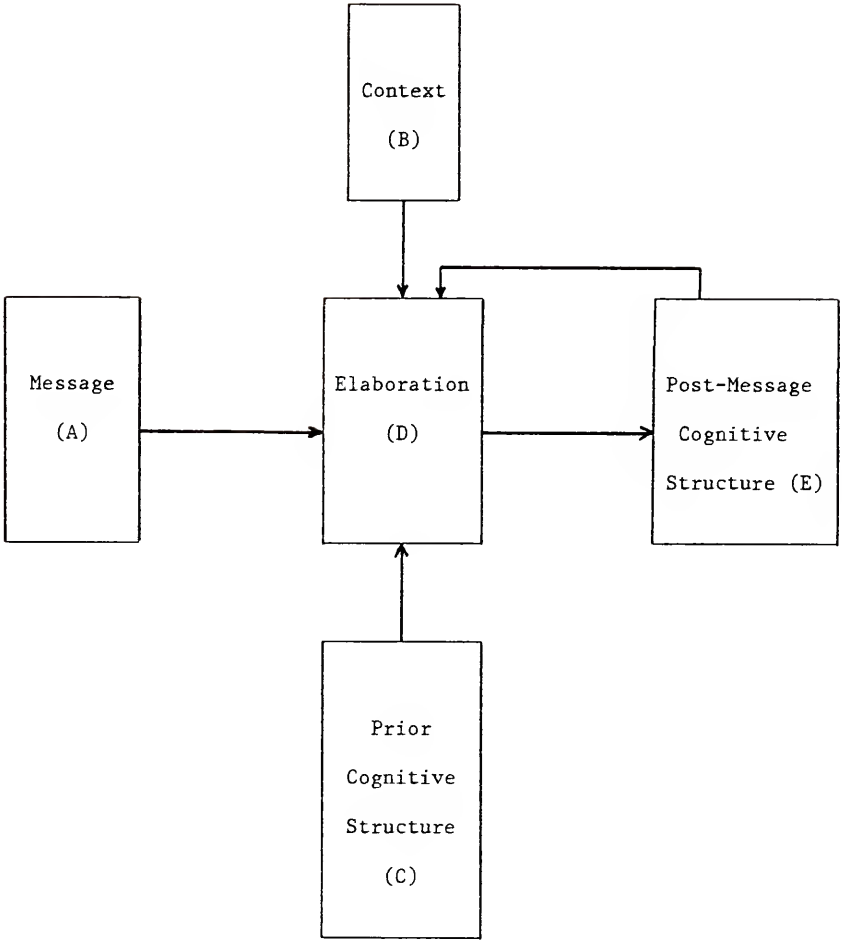


Figure 5. An Elaboration Model of Message Processing

persuasiveness of the arguments used. A final influence on elaboration is the post cognitive structure--changes induced early in the episode affect the manner in which following portions of the message are elaborated.

Relationships Portrayed by the Model

Given the model presented in Figure 5, many relationships may be selected for investigation: the effect of message variables on the post cognitive structure or on elaboration (A-E or A-D relationships); the effect of context on elaboration or on the post cognitive structure (B-D or B-E relationships); the effect of pre-existing cognitive structure on those same two variables (C-D and C-E relationships); and finally, the effect of elaboration on the post cognitive structure (D-E relationships). Interactive effects of context, message, and pre-existing cognitive structure on elaboration or on the modified cognitive structure may also be examined.

The remainder of this chapter will examine studies of persuasion that have presumed that subvocal responses are an integral part of the persuasion process--that subvocal responses play an intervening role in the processing of persuasive communications. This discussion will begin with an examination of the predictor variables used in these studies vis-a-vis their relationship to the elaboration model. The operationalization of these predictors, as well as their proposed role in the persuasion process, will be examined.

Next will follow a review of the criterion variables that have been used in these studies. These criterion measures will also be positioned relative to the elaboration model. The operationalization of the

criterion variables will be examined for their appropriateness and their validity.

The third part of this examination of persuasion studies will focus on their findings. This discussion will proceed by examining, in turn, each relationship which is portrayed in the elaboration model. It is organized by the type of criterion measure which is used. Studies of similar relationships are reviewed for convergent or divergent findings.

Predictor Variables

Distraction:

A very frequently studied relationship is that of distracting context variables on either elaboration or on the post cognitive structure (B-D and B-E studies). Distraction studies have at times also included message and/or pre-existing cognitive structure variables so that interactive effects could be examined.

The proposed role of distraction in processing persuasive communications:

Even the first distraction studies, which studied the effect of contextual variables on the post cognitive structure assumed that elaboration was an intervening variable in the relationship.

Festinger and Maccoby suggested this scenario for a listener who was confronted with a counterattitudinal message: "...It is most likely that under such circumstances, while he is listening to the persuasive communication, he is very actively, inside his own mind, counterarguing, derogating the points the communicator makes, and derogating the communicator himself" (1964, p. 360).

The presumption was that if negative elaboration could be prevented by distraction, then persuasion would be enhanced. Most of the researchers of the distraction effect have presumed that elaborative processing can only proceed after successful reception and comprehension (processing to the surface structure level). If this were true, then distraction could interfere with additional processing (including negative elaboration), thus leaving the message relatively intact within the recipient's cognitive structure. "As the contiguity model implies, the distractor must not be so strong as to interfere with understanding or learning of the persuasive message, but it must be strong enough to interfere with production of counters" (Maccoby and Roberts 1972, p. 262). Bither also suggests this sequential processing, noting that counterarguing ("supposedly inhibited by distraction..." (1972, p. 1)) is primarily involved in the yielding stage, but may also affect reception. He concludes that "...the range of distraction expected to inhibit counterarguing and not completely inhibit message reception is potentially quite narrow" (1972, p. 1). However, it seems that a person who could not fully elaborate a message because of the competing demands of a distractor, could elaborate by using non-message cues such as the source.

The elaboration model suggests that the relationship between distraction and persuasion may be curvilinear. As distraction increases its demands on the limited processing capacity, the initial effect may be that responses to propositions offered by the communicator may be interrupted. If these responses would have countered the communicator's propositions, then the recipient's representation of the message may lack these counters. However, at higher levels of distraction, reception

of the propositions would be affected, and these propositions would not be included in the recipient's resultant, propositional representation of the communication as accommodated with his/her cognative structure. Wyer (1974, Ch. 7) suggests such an effect based upon McGuire's (1968) theory of persuasibility. However, this curvilinear prediction rests upon the assumption that elaboration must follow reception of communication content.

Given that the recipient may be aware of the content of the communication, a recipient who is highly motivated to resist persuasion may use strategies of resistence which require less of the limited processing capacity than do point-by-point, content-based counter-argumentation and which may proceed without (or intermittently with) registration of message content. (Wright 1974b suggests that subjects may switch to responses directed toward the source when time for responding is restricted).

Earlier, Kelman (1953) attempted to reduce negative elaboration by manipulating motivation. He assumed that responses, which he thought would usually be "self-verbalizations" (1953, p. 187), could be either interfering or supporting. For example, he defined interferring responses as "...any implicit response made by the individual which provides motivation against the overt response he makes; which limits the stimulus situations to which the overt response is applicable; or which is generally irrelevant (such as aggressive or distracting responses)" (1953, p. 187). His basic hypothesis sounds almost identical to that advanced in later distraction studies, yet pre-dates them by eleven years: "The performance of a response in the communication situation will facilitate transfer, and hence increase attitude change, to the

extent to which implicit supporting responses are produced; it will impede transfer and hence decrease attitude change, to the extent to which implicit interfering responses are produced" (1953, p. 187-188). Although he did not measure these implicit responses directly, he did elicit a subjective report of the amount of thinking during the task. The results supported his hypothesis.

Alternative explanations of distraction effects:

The effects of distraction studies have been explained in ways other than postulating interference with elaboration. Baron, Baron, and Miller (1973) elaborate these explanations in an excellent review article.

One explanation is the Affect Hypothesis (Baron, Baron, and Miller 1973; p. 313; see also McGuire 1966, pp.481-482). This hypothesis suggests that subjects enjoy the distractor and, therefore, feel more positively toward the position advocated in the message. "In sum, it is clear that while the affect hypothesis may have some validity it is inadequate as a single explanation for the relationship between distraction and persuasion" (Baron, Baron, and Miller 1973, p. 314).

A second explanation suggests that distraction effects will only be found in reactive settings. The effect could occur two ways: the Ss could be unaware of the persuasive intent (Rosenblatt 1966); or Ss may perceive the study as an attempt to evaluate their ability to process the message in the presence of the distractor. "...Subjects motivated by evaluation apprehension would have marked the attitude questionnaire in the direction of agreement with the message, not on the basis of actual attitude change, but to demonstrate to the experimenter their powers of concentration" (Silverman and Regula 1968, p. 274).

The *Effort Hypothesis* is a third alternative to thought disruption as an explanation of distraction effects: "...the effort required to comprehend a counterattitudinal message will directly determine the amount of dissonance created by the choice" (Baron, Baron, and Miller 1973, p. 317). The effort expended in listening to a message in the presence of a distractor could be justified "...by either overvaluing the message or distorting their initial attitude so as to minimize any discrepancy with the message" (Baron, Baron, and Miller 1973, p. 317).

Another explanation considers *learning the content* of the message. "From a simple learning theory interpretation one would expect distraction to have the opposite effect [from that proposed in the thought-disruption hypothesis], interfering with comprehension of the argument and thus lowering persuasive impact" (McGuire 1966, p. 481). This argument has led some researchers to limit predictions based on the thought disruptions hypothesis: "...when they [distractors] interfere with counterarguing without substantially blocking message reception, persuasibility is likely to be increased" (Bither 1972, p. 1). "This may suggest that distraction serves to facilitate acceptance of a counterattitudinal message only when the distraction is not so severe as to inhibit reception of the arguments contained in the message" (Osterhouse and Brock 1970, p. 355).

Kiesler and Mathog (1968) suggest that distractors may have a negative effect on persuasion, since they may *annoy* Ss who are trying to listen to a communication.

Conceptualizing and operationalizing distraction:

The generalization of findings from distraction studies is hampered by the many, very different conceptual and operational definitions used

for distraction and by researchers' failure to distinguish between the distractor and distraction. A distractor generally refers to some stimulus which induces a response (distraction) that restricts the extent to which a person can process a message. It is assumed that the situation requires a division of attention between the processing of the message and of the induced response.

The definition of distraction that will be used for the remaining discussion relies upon the concept of a limited processing capacity presented by Norman and Bobrow (1975). Figure 6 again presents their model which was discussed in Chapter 2. In this figure, $\underline{L}$ represents the maximum processing capacity (attention) available for allocation among various concurrent tasks at any given time. $\underline{P}_A$ represents the amount of elaboration that could occur if all of the available attentional capacity were devoted to elaboration. The performance curve for processing the distractor is depicted in this figure as one for a task which requires relatively less processing capacity to perform well than does elaboration. Such a process is less conceptually driven and more data driven than is elaboration. (A data driven task requires simple responses to environmental stimuli, while a conceptually driven task requires more extensive processing--utilizing expectations and establishing new linkages between schemata. Signal detection would be a data driven task, while elaboration would be a conceptually driven task.) $\underline{R}_{distractor}$ represents the minimum capacity that must be allocated to the distractor to process it in an error-free fashion. If this much capacity were allocated to the distractor, then only the remainder, $\underline{R}_{elaboration}$ would be left for possible allocation to elaboration. If this remainder were allocated to elaboration then performance on

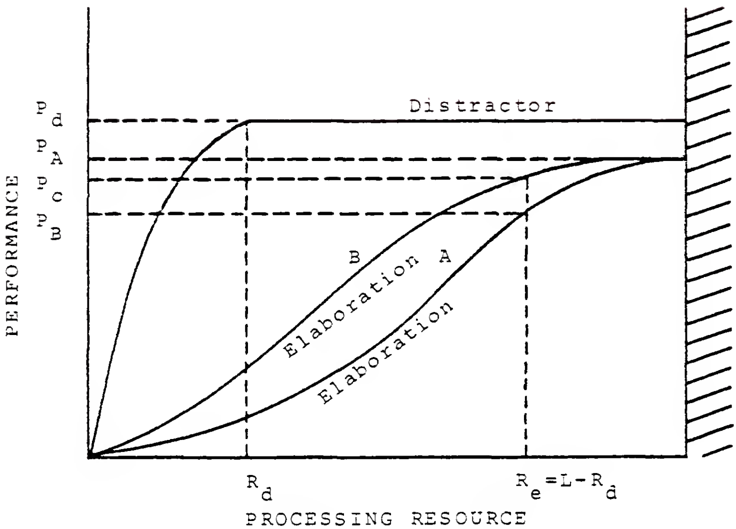


Figure 6. The relative allocation of the processing resource and the resultant performace curves when attending to the distractor is perceived as the primary task: showing the changes anticipated with increasing familiarity of the message. (After Norman and Bobrow).

elaboration would be reduced from $\underline{P}_A$ to $\underline{P}_B$. Allocation of any resource level below $\underline{R}_{distractor}$ would allow a higher level of performance on elaboration (provided that the increase in processing resource were allocated to elaboration).

If the message were more familiar to the recipient, the performance-resource curve for elaboration would shift to the left (such as from A to B in Figure 6) as elaboration would then require less attentional capacity (see Norman and Bobrow 1975, p. 61). With such a shift, the decline in elaboration due to distraction would be reduced and a performance level on elaboration such as P_c would be expected.

If a person viewed attending to the message as the primary task and sufficient attention were devoted to the message to fully elaborate it, then performance on the distracting task should suffer (provided that the remaining attentional capacity was insufficient for error-free performance on the distracting task). This situation is depicted in Figure 7.

Zimbardo et al. (1970) explicitly considered this issue. "Since the variety of ways in which distraction has been manipulated is almost equal to the number of studies performed, and if different types of distraction activities elicit different amounts of attention to them, the failure to explicitly control the focus of the subject's attention could account for the current diversity of the findings" (1970, p. 671). To control for this, they developed instructions to engender either a distraction-task set or a message set in their subjects.

However, it is important to remember that processing requires both time *and* space. As such, distraction may result from distracting stimuli that limit time *or* space *or* both. *Space* is best understood as

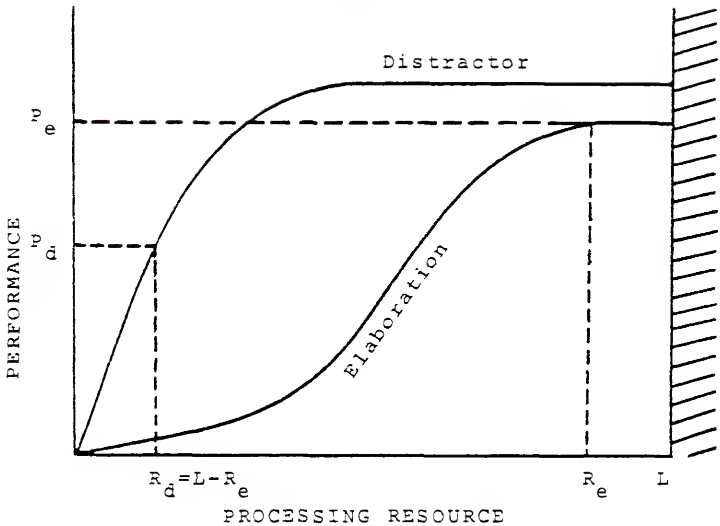


Figure 7. Relative allocation of the processing resource and the resultant performance curves when elaboration is perceived as the primary task. (After Norman and Bobrow 1975).

the limited processing capacity of an individual (Norman and Bobrow 1975). In this case, a distractor would be a stimulus which elicits a response that competes <u>concurrently</u> for the processing resource that is needed for elaboration of a message.

Intermittent distractors pose a different problem from concurrent distractors for the message processor. Here the effect may be more disruptive than competing and the recipient may be able to sequentially process both the distractor and the message. The ability to do this should increase with the predictability of the message (due to familiarity or to redundancy within the message). The recipient could retain the message at one level of processing while attending to the distractor, and process it beyond that level when elaboration resumes--yet the overall <u>time</u> available for elaboration will be reduced. The more predictable the occurrence of the distractor, the more successful should be the sequential processing, since message processing could be paced in anticipation of the distractor.

Other operationalizations of distraction do not inherently require time or space for processing, but instead affect message <u>signal quality</u> (e. g. static--Silverman and Regula 1968; reduced video quality--Keating and Latané 1972; and foreign accent--Bither 1972 and Bither and Wright 1973). It is uncertain that these manipulations qualify as distractors, since they require no response on the part of the recipient. (See Osterhouse and Brock 1970, p. 355.) In fact, variations in signal quality may yield a <u>greater</u> attention allocation for message processing because of attention <u>attracting</u> properties and because motivated Ss may allocate more attention to the message in an effort to process it successfully.

A conceptual definition of distraction. For the remaining discussion, distraction will be defined as the allocation of a portion of a message recipient's processing capacity to implicit (cognitive) or overt responses, either continuously or intermittently during message reception. This is similar to the definition used by Insko, Turnbull, and Yandell (1974, p. 508) "...distraction involves a division of attention between a persuasive communication and communication irrelevant stimuli."

The definition proffered here differs from others that stipulate the amount of the limited processing capacity allocated to the distractor. For example, Keating and Brock (1974, p. 308) define distraction this way: "By definition, a distraction is not the 'center of attention.' Distraction is in competition for the main focus of attention, but does not absorb this attention." This definition probably reflects wishful thinking on the part of the experimenters. In fact, the message recipient may view the message as the distractor and the intended distractor may be the "center of attention" as far as the subject is concerned. Zimbardo et al. 1970 manipulated the perceived primary task--the message or the distractor. (Their manipulations and findings will be discussed in later sections.)

Another definition of distraction that has been used presumes knowledge of both B-D (context-elaboration) and D-E (elaboration-post cognitive structure) relationships. This definition makes the presumed effects of distraction a part of the definition: "If attitude change is more apt to be induced due to interference with counterargument, then this defined as distraction" (Gardner 1970, p. 26).

Construct valididy of distraction manipulations. Since distraction involves a response on the part of a message recipient, the presence of a potential distractor does not guarantee that distraction will occur. Therefore, conclusions and generalizations drawn from distraction studies require evidence that the distraction manipulations were successful.

There are three levels of generalization at issue here. The first concerns whether, in any given distraction study, variations in a criterion variable may be attributed to varying levels of distraction, i.e., internal validity. For a distraction study to have internal validity, it must be known that differing amounts of the limited processing capacity were allocated to the distracting task in the various conditions _and_ that no other variables were confounded with the differential allocation of attention, i.e., no other variables should change concurrently with level of distraction. For internal validity, knowledge of ordinal gradations in attention allocation (free of confounding) is sufficient to attribute changes in the criterion to changes in distraction. However, an ordinal ranking of distraction conditions is insufficient for a thorough understanding of the phenomenon, especially if the effects should vary curvilinearly with increasing distraction. To generalize the nature (i.e., linear or curvilinear) as well as the existence of changes in the criterion to changes in distraction requires knowledge of absolute, as well as relative, differences in attention allocation across distraction conditions.

A second generalization issue deals with inter-study comparisons--external validity. Here, again, knowledge of absolute levels of

traction are necessary. In the absence of quantitative knowledge of levels of distraction that have been achieved in various studies, verification through replication cannot be effected. Even though different researchers may concur in labeling their conditions (e.g., low, moderate, high), the actual levels which they have achieved may be quite different. They may have achieved levels of distraction which engage very different levels of the limited processing capacity. If the effects of distraction are curvilinear, then seemingly contradictory findings may be consistent, if it were shown that all of the levels of distraction in one study were low and those of the other were high. The inferred inflection point could, then, correspond to some intermediate level of distraction (the possible level in some future study).

The third generalization issue deals with the prediction that effects observed in the experimental setting will also occur in a natural setting--ecological validity. The primary problem here arises from the flexibility and the partly voluntary nature of the attention allocation process. Without the involvement on the part of the recipient and the measure of situational control exercised by the experimenter, a distractor which had an impact in the experimental setting may simply be ignored in the natural setting.

Table 1 shows the dimensions that have varied in distraction manipulations of 22 studies. Clearly, more than the level of distraction was manipulated in these studies. Each of these dimensions will be explained and their implications for the generalizability of the studies' findings will be discussed.

Distraction mode pertains to the manner in which the distractor is presented. This is significant for both internal and external validity.

Table 1. Components of Distraction Settings.

Authors and Year	Distractor	Distractor Mode	Distractor Expected?	Instructed to Attend to Distractor?	Intermittent or Continuous Distractor?
Festinger and Maccoby 1964	humorous film	visual	no	no	continuous
Freedman and Sears 1965	personality of speaker	visual	yes	in one condition	continuous
Rosenblatt 1966	irrelevant slides	visual	yes	no	continuous
Gardner 1966	slot car operation	visual	yes	yes	continuous
Vohs and Garrett 1968	1. geometric figure task 2. arithmetic	1. print 2. print	1. yes 2. yes	1. yes 2. yes	1. ? 2. ?
Kiesler and Mathog 1968	number copying task	print	yes	yes	?
Haaland and Venkatesan 1968	1. multiple-choice and semantic differential questionnaire 2. humorous film	1. print 2. visual	1. yes 2. no	1. yes 2. no	1. ? 2. continuous
Venkatesan and Haaland 1968	1. multiple-choice and semantic differential questionnaire 2. unrelated video	1. print 2. visual	1. yes 2. no	1. yes 2. no	1. ? 2. continuous
Silverman and Regula 1968	static	audio	yes	no	continuous
Baron and Miller 1969	personality of speaker	N.A.	yes	yes	N.A.
Osterhouse and Brock 1970	flashing lights	visual	yes	yes	intermittent (varying interval)
Rule and Rehill 1970	22 different sound effects	audio	no	no	continuous
Gardner 1970	same as Gardner (1966)				
Zimbardo et al. 1970	number summation task	print	yes	yes	?
Keating and Latane 1972	reduced video quality: 1. continuous 2. intermittent	1. visual 2. visual	1. no 2. no	1. no 2. no	1. continuous 2. intermittent (fixed interval)
Bither 1972	1. foreign accent 2. views of umbrellas (color) 3. slow motion football (action (B & W)	1. audio 2. visual 3. visual	1. no 2. no 3. no	1. no 2. no 3. no	1. continuous 2. continuous 3. continuous
Bither and Wright 1973	same as Bither (1972)				

(continued)

Table 1--extended.

Voluntary Allocation of Attention to Distractor?	Self or External Pacing?	Overt Response to Distractor Required?	Data or Resource Limited Response?	Distractor Response Mode	Message Mode	Message Expected?	Message Length	Instructed to Attend to Message?
yes	external	no	?	?	audio (film)	yes	10 min.	yes
yes	external	no	?	?	audio (person)	yes	12 min.	in one condition
yes	external	no	?	?	audio (tape)	yes	?	yes
no	external	yes	data	motor	audio (tape)	no	?	no
1. yes 2. yes	1. self 2. self	1. yes 2. yes	1. resource 2. resource	1. written* 2. written*	audio	yes	?	yes
yes	self	yes	data	written*	audio	yes	?	yes
1. yes 2. yes	1. self 2. external	1. yes 2. no	1. resource 2. ?	1. written* 2. ?	audio	yes	10 min.	yes
1. yes 2. yes	1. self 2. external	1. yes 2. no	1. resource 2. ?	1. written* 2. ?	audio	no	1 min.	no
yes	external	no	?	?	audio	yes	8 min.	yes
N.A.	N.A.	no	N.A.	?	audio (anticipated)	yes	?	in one condition
no	external	yes	data	vocal	audio	yes	6 min.	yes
yes	external	no	?	?	print	yes	6 min. limit	yes
yes	self	yes	resource	written*	audio	yes	5 min.	yes
1. yes 2. yes	1. external 2. external	1. no 2. no	1. ? 2. ?	1.? 2.?	audio	yes	3½ min.	yes
1. yes 2. yes 3. yes	1. external 2. external 3. external	1. no 2. no 3. no	1. ? 2. ? 3. ?	1. ? 2. ? 3. ?	audio	yes	1 min.	yes

Table 1--continued.

Authors and Year	Distractor	Distractor Mode	Distractor Expected?	Instructed to Attend to Distractor?	Intermittent or Continuous Distractor?
Keating and Brock 1974	flashing lights: 1. visual monitoring 2. vocal monitoring 3. manual monitoring 4. vocal and manual monitoring	1. visual 2. visual 3. visual 4. visual	1. yes 2. yes 3. yes 4. yes	1. yes 2. yes 3. yes 4. yes	1. intermittent (varying intervals) 2. intermittent (varying intervals) 3. intermittent (varying intervals) 4. intermittent (varying intervals)
Insko et al. 1974	same as Zimbardo et al. (1970)				
Petty, Wells and Brock 1976	X's flashed on a screen	visual	yes	yes	intermittent (fixed interval)
Maccann Haslett 1976	visual or auditory stimulus: 1. visual monitoring 2. vocal monitoring 3. written monitoring	visual or audio	1. yes 2. yes 3. yes	1. ? 2. yes 3. yes	1. intermittent (fixed intervals) 2. intermittent (fixed intervals) 3. intermittent (fixed intervals)
Petty 1977	same as Petty, Wells and Brock (1976)				

* subvocalizations may accompany these responses

Table 1--continued--extended.

Voluntary Allocation of Attention to Distractor?	Self or External Pacing?	Overt Response to Distractor Required?	Data or Resource Limited Response?	Distractor Response Mode	Message Mode	Message Expected?	Message Length	Instructed to Attend to Message?
					audio	yes	6 min. &	?
1. yes	1. external	1. no	1. ?	1.?				
2. no	2. external	2. yes	2. data	2. vocal				
3. no	3. external	3. yes	3. data	3. manual*				
4. no	4. external	4. yes	4. data	4. vocal & manual				
no	external	yes	data	written*	audio	yes	3 min.	yes
					audio	yes	7 min.	?
1. yes	1. external	1. no	1. data	1. ?				
2. no	2. external	2. yes	2. data	2. vocal				
3. no	3. external	3. yes	3. data	3. written*				

Different levels of distraction within any one study should not be confounded with mode of presentation (e.g., high distraction/visual and low distraction/audio). If confounding exists, then an equally good case can be made for either variable as the cause of any observed effects. In addition, studies which utilize different modes of presentation are not strictly comparable--especially in the absence of an independent, quantitative measure of distraction.

Whether or not the distractor is expected is another significant dimension. When Ss have not been instructed that the distractor will appear, its sudden appearance could have at least two undesirable effects. It could startle the Ss (possibly desirable if the effect of a startling distractor is being studied) and it could lead Ss to form the hypothesis that the experimenter is testing their ability to perform the primary task in a difficult situation. Depending on their reaction to this hypothesis, they may try to ignore the distractor to impress the experimenter or they may exhibit a reactance effect and purposefully do badly on the primary task. If these demand effects occur, then they pose alternative explanations for any observed effects, whenever a no-distraction condition exists. When the unexpectedness of the distractor varies between studies, differential observed effects may be a result of this factor.

The next factor considered in Table 1 is whether Ss were instructed to attend to the distractor. When Ss are instructed to attend to the distractor, it is no longer left to chance whether Ss will perceive the distractor to be a salient aspect of the experiment. However, in these studies demand effects will always be a potential alternative explanation for differences observed between distraction and no-distraction

conditions. Differences between studies that did and did not include instructions to attent to the distractor may be due to this factor. The generalization of findings from studies in which Ss were instructed to attend to the distractor to natural settings is very tenuous.

Instructing Ss to attend to a distractor does not ensure that they will do so. Another factor that has varied both within and across distraction studies is whether allocation of attention to the distractor is voluntary or whether the task requires that some attention be allocated to the distractor. In studies in which allocation of attention is voluntary it is difficult to know if distraction has actually occurred.

This problem is especially important in light of the fact that allocation of processing resources to a distractor in natural settings is largely voluntary, as was seen in the previous chapter in the discussion of the "cocktail party problem". With the exception of stimuli which automatically activate a particular schema (Schneider and Shiffrin 1977, p. 2), humans have the ability to select which stimulus from among many competing ones they will process. This ability greatly reduces the generalizability of distraction findings from studies in which allocation of processing capacity is non-voluntary to any planned effort to exploit the effect in natural exposure condition, since people may not attend to the distractor under those conditions. However, the findings may help in explaining communications' effects resulting from uncontrolled distracting stimuli accompanying message reception. Table 1 indicates that allocation was voluntary in 15 of 22 studies reviewed.

The next factor, continuity of the distractor is significant because of the different problems confronting Ss in a condition in which

there is a continuous distractor and those in a condition with an intermittent distractor. Ss facing an intermittent distractor may be able to alternately process the distractor and the message while those facing a continuous distractor would not be able to utilize such a switching strategy without missing some of the message or without a decrement in performance on the distracting task.

Intermittent distractors may be either fixed interval or varying interval. Fixed-interval, intermittent distractors should be easier to cope with than varying-interval, intermittent distractors, as Ss may learn to correctly anticipate the occurrence of the distractor and make compensatory changes in message processing. Therefore, the effect of fixed-interval, intermittent distractors (e.g., Keating and Latané 1972; Petty, Wells, and Brock 1976; Maccann Haslett 1976; and Petty 1977) should be less that that of varying-interval, intermittent dis tractors (e.g., Osterhouse and Brock 1970; Keating and Brock 1974).

As with the other factors, if different levels of distraction within a study vary with continuity, then continuity is an alternative explanation for any observed effects.

The rate of occurrence of a distractor in a distraction study may be either self-paced by the subject or externally-paced by the experimenter. With self pacing, the message recipient should be able to intermittently process the message and the distractor by relying on natural breaks in the communication, redundancy, grammar and familiar schemas for that type of communication to provide the time necessary to process the distraction The subjects also controls the ultimate frequency and, hence, the number of intrusions of the distractor during message reception.

When this locus of control (i.e., external or internal) is confounded with level of the distraction as in the Haaland and Venkatesan (1968) study, it may be the cause of observed effects on the criterion. Studies in which self-pacing is allowed are not directly comparable to those in which the pace is determined externally. The rate of intrusion of distractors in natural settings is probably most frequently self-paced so that findings from studies in which the distractor was externally paced may have little ecological validity.

For some manipulations of distraction, an overt (behavioral, externally visible) response was required of the Ss. Studies in which no overt response was required leave unanswered the question of whether a response (and thereby distraction) occurred. Any such manipulation would require a manipulation check (to be discussed in a following section) to establish that distraction had been manipulated (a prerequisite for internal validity). Manipulation checks could also allow inter-study comparisons, provided that the measures of distraction in the various studies are comparable. Many naturally occurring distractors also require no overt response. However, the manipulation checks are still necessary for a study to have ecological validity, since it is the only way of knowing exactly what effect is to be generalized. Table 1 indicated that no-overt-response conditions existed in 12 of 22 studies reviewed.

Whether the response to the distractor is data-limited (data-driven) or resource-limited (conceptually driven) determines the amount of the limited processing capacity which will remain for message processing if the distractor is processed fully (see Figure 5). Conceptually driven processing of a distractor should require more of the

limited processing capacity than data-driven processing and should have a greater adverse effect on elaboration.

If the full range of distraction is to be represented in a study, then the data/resource limited aspect of the distraction response is a natural confound with level of distraction. However, this aspect should be explicitly considered so that it will be known to what range of distraction the effects are to be attributed within a study. The findings of studies which utilize levels of distraction occurring within different ranges may seem contradictory unless the data/resource limited nature of the response is considered.

The distraction response mode is significiant because two tasks with the same mode interfere more with each other than two tasks which utilize different modes (Posner 1973, p. 144; Brooks 1968, p.354). For studies in which modality of response is confounded with level of distraction (e.g., Haaland and Venkatesan 1968; Bither 1972; and Bither and Wright 1973) modality of response is an alternative explanation for the findings. Readers should be wary of responses to distractors which are presented as written or manual, as it may be very difficult for a person to perform a written or manual response without accompanying subvocalizations.

In comparing between studies, both the response mode and message mode should be considered as it is likely that they interact. Which modes are involved also affect a study's ecological validity. For example, an audio distractor may have one effect on a visually presented message in the natural setting and another effect on an audio message.

The same distraction-response modality has been interpreted (perhaps justifiably) as representing the manipulation of different

conceptual variables in two separate studies. Keating and Brock (1974) considered "manual monitoring" and "vocal monitoring" of a distracting stimulus as two distinct conditions, while Maccann Haslett (1976) considered these same treatments collectively as "low active." The "vocal-and-manual" monitoring task in the Keating and Brock (1974) study was considered "high active" by Maccann Haslett (1976). These are obviously two legitimate ways of viewing the same distractor--one is concerned with the modality of distraction and the other with overall activity level. Unfortunately, the modality was not established in either study.

Qualitative vs quantitative operationalizations of distraction. In order to generalize the findings of distraction studies to non-experimental settings, it is necessary that the dimensionality of the distractor and of the distraction be known. (Table 2 presents the hypothetical dimensions manipulated in the studies reviewed here.)

When the distractor varies in some qualitative way, it is often difficult to establish just what dimension is involved. When the dimensionality is not known, the study degenerates into a copy-test situation in which the findings can only be generalized to the versions of the distractor present in the study. When dimensionality is not known, attribution of causal effects to a particular variable is subject to whimsical labeling by the experimenter. The result is multiple, alternative explanations of the observed phenomenon.

For example, in the Festinger and Maccoby study (1964), does the distractor represent a point on the humor dimension, on an irrelevancy dimension, on the attractiveness of the actors, or on some other dimension? What is the dimensionality of paying attention to the personality

Table 2. Distraction Manipulation Checks.

Authors and Year	Hypothetical Dimension	Condition Number	Levels of the Predictor Variable	Manipulation Check Measure	Manipulation Check Means by Condition	Significance
Festinger and Maccoby 1964	?	1 2	film of speaker humorous film	none	(1) (2) --- ---	---
Freedman and Sears 1965	?	1 2	content set personality of the speaker	none	(1) (2) --- ---	---
Rosenblatt 1966	oddity of slides	1 2 3 4	no slides "moderate" (dental hygiene) "strong" (dental hygiene) "strong" (general psych. course slides)	none	(1) (2) (3) (4) --- --- --- ---	---
Gardner 1966	divided attention	1 2 3	undivided one slot-car (slow) two slot-cars (fast)	"amount of perceived effort and difficulty"	(1) (2) (3) values not reported	---
Vohs and Garrett 1968	complexity	1 2 3	listen only geometric figure task arithmetic	none	(1) (2) (3) --- --- ---	---
Kiesler and Mathog 1968	?	1 2	listen only ("low interference") number copying task ("high interference")	"how difficult was it to concentrate?"	(1) (2) values not reported	---
Haaland and Venkatesan 1968	visual vs behavioral	1 2 3 4	normal videotape of speaker humorous film multiple choice and semantic differential questionnaire humorous film and questionnaire	none	(1) (2) (3) (4) --- --- --- ---	---

(continued)

Table 2 - continued.

Authors and Year	Hypothetical Dimension	Condition Number	Levels of the Predictor Variable	Manipulation Check Measure	Manipulation Check Means by Condition	Significance
Venkatesan and Haaland 1968	visual vs behavioral	1	normal commercial	none	(1) ---	---
		2	"unrelated" video		(2) ---	
		3	multiple choice and semantic differential		(3) ---	
		4	unrelated video and questionnaire		(4) ---	
Silverman and Regula 1968	?	1	no static	none	(1) ---	---
		2	"low" static		(2) ---	
		3	"high" static		(3) ---	
Baron and Miller 1969	?	1	content set	none	(1) ---	---
		2	personality of the speaker		(2) ---	
Osterhouse and Brock 1970	rate of presentation	1	listen only	"To what extent do you feel that you were distracted while listening to the speech?" (70 point scale)	(1) 11.3	$p < 0.001$
		2	10 flashes/min.		(2) 27.9	
		3	20 flashes/min.		(3) 38.8	
		4	30 flashes/min.		(4) 44.7	
Rule and Rehill 1970	?	1	read only	change in heartrate	(1) +4.71 beats/min.	$p < 0.05$
		2	22 different sound effects		(2) +8.22 beats/min.	
Gardner 1970	same as Gardner 1966 (above)					
Zimbardo et al. 1970	perceived primary task	1	message only	1. "Perceived effort expended in listening to and comprehending the message" (100 pt. scale)................	(1) 70 (2) 59 (3) 28	$p < 0.01$
		2	message set (concurrent number summation)	2. Percent of attention allocated to the communication (100 pt. scale)................	(1) 81 (2) 65 (3) 45	$p < 0.01$
		3	task set (concurrent message)	3. Change in performance on number task....................	(1) --- (2) -14.5 (3) +19.3	$p < 0.001$
Keating and Latane 1972	continuous vs intermittent	1	undistorted TV	none	(1) ---	---
		2	continuous distortion		(2) ---	
		3	intermittent distortion		(3) ---	

(continued)

Table 2 - continued.

Authors and Year	Hypothetical Dimension	Condition Number	Levels of the Predictor Variable	Manipulation Check Measure	Manipulation Check Means by Condition	Significance
Bither 1972	audio vs visual and strength of visual	1 2 3 4	video scenes related to the message "audio" (French accent) color views of umbrellas "mild visual" B & W slow-motion football "strong visual"	sum of two scales: "degree of attention focused on the message and the amount of the message that was understood" **	(1) (2) (3) (4) 1.8 2.6 2.38 3.08 ("audio" and "mild visual" not significantly different)	$p < 0.001$
Bither and Wright 1973	same as Bither (1972)	1-4	same as Bither 1972	sum of two scales: "degree to which the person was able to attend to and to understand the ad's message" **	(1) (2) (3) (4) 4.91 7.54 6.90 8.91 "significant differences between the three visual distraction treatments"	p value not reported
Keating and Brock 1974	vocal vs non-vocal and rate of presentation	1* 2* 3* 4* 5* 6* 7*	no flashes 10 flashes/min. 25 flashes/min. visual monitoring manual monitoring vocal monitoring vocal and manual monitoring	"the extent to which they felt distracted" (70 pt. scale)	(1) (2) (3) 15.9 33.2 42.2 (4) (5) (6) (7) values not reported	$p < 0.001$ p value not reported
Insko et al. 1974	same as Zimbardo (1970)	1 2 3	message only message set (concurrent number task set (concurrent message)	1. "effort exerted listening to the communication" (100 pt. scale)...................... 2. "division of attention between the two tasks" (100=all message")..................	(1) (2) (3) 63.32 61.69 44.09 --- 51.87 22.11	$p < 0.01$ $p < 0.01$

(continued)

Table 2 - continued.

Authors and Year	Hypothetical Dimension	Condition Number	Levels of the Predictor Variable	Manipulation Check Measure	Manipulation Check Means by Condition	Significance
Petty, Wells and Brock 1976 (experiment 1)	frequency	1 2 3 4	no flashes 15 sec. intervals 5 sec. intervals 3 sec. intervals	1. "To what extent did you find the task effortful?" (7 pt. scale, 1 = completely effortful)....................	(1) (2) (3) (4) 4.62 4.80 3.68 2.64	$p < 0.001$
				2. Distracted "...from paying attention to the message" (7 pt. scale, 7 = completely distracted)........	2.21 3.30 4.61 4.81	$p < 0.001$
(experiment 2)	frequency	1 2	15 sec. intervals 5 sec. intervals	1. "effortful" (scale categories not reported)......	5.00 6.70	$p < 0.05$
				2. "distracted" (scale categories not reported)......	3.93 6.93	$p < 0.005$
Maccaun Haslett 1976	activity and intensity	1* 2* 3* 4* 5* 6*	no numbers presented 10/min. 22/min. "passive" (visual monitoring) "low active" (vocal or manual monitoring "high active" (vocal and manual monitoring)	none	(1) (2) (3) --- --- --- (4) (5) (6) --- --- ---	--- ---
Petty 1977	same experiment reported in Petty, Wells and Brock (1976)					

* Since the experiment involves a factorial design, these are not legitimate conditions, but rather are levels of the respective factors.
** It is likely that these are both paraphrases of the same questions.

of the speaker (Freedman and Sears 1965; Baron and Miller 1969)? How can "oddity" of slides (Rosenblatt 1966) be generalized to another experimental study, much less to non-experimental situations? Why is color views of umbrellas "mild visual" and black-and-white, slow-motion football action "strong-visual" (Bither 1972; and Bither and Wright 1973)? Any observed effects may only apply to umbrellas and football and not to distractors in general.

Quantitative manipulations of distractors do not suffer from the same ambiguity. However, quantitative distractors require that the distractor be externally paced and that allocation of processing capacity be non-voluntary. These requirements, make generalization difficult to non-experimental situations, as that is exactly where allocation of attention is voluntary and self-paced.

Requirements of manipulation checks of distraction. The definition of distraction which is being used here provides criteria for evaluating the manipulation of distraction in communications studies. Distraction can be said to have been manipulated if one of the following conditions is met: a) differing amounts of attentional capacity were allocated to a continuous distractor across distraction conditions; or b) the frequency at which the allocation of attention is elicited by the distractor varies across distraction conditions. Both of these conditions require knowledge of whether or not attention has been allocated to the distractor.

Establishing the construct validity of a distraction manipulation requires more than showing that distinct levels of some measure of the distraction have been achieved--it must also be shown that the purported measure of the distraction is indeed a measure of distraction. The

measure should establish the amount of processing capacity that is allocated to processing the distractor. This is always true for continuous distractors and is necessary for generalizing the effects of intermittent distractors and comparing different studies which utilize intermittent distractors.

Applying the criteria to distraction studies. Table 2 presents the levels of distraction which were ostensibly achieved in 22 studies. The studies which seem to meet the criteria for distraction manipulation and which should have interval validity without the use of a manipulation check are those which manipulated frequency of presentation of the distractor, required a response from the subjects, and for which allocation of attention to the distractor was non-voluntary: Osterhouse and Brock 1970; Keating and Brock 1974--conditions 2-4; Petty, Wells and Brock 1976; Maccann Haslett 1976--conditions 2 and 3; and Petty 1977. In these studies, knowledge of the Ss' successful completion of the task is synonymous with knowledge of a successful manipulation. For other manipulations it is necessary to have manipulation checks to establish whether different levels of distraction were actually achieved.

A review of Table 2 reveals the problem of generalizing findings across studies (external validity), as well as the problem of internal validity given the arbitrary nature of the labeling of distraction conditions. For example, does attending to the personality of a speaker require more or less processing capacity than attending to a humorous film or watching dental hygiene slides (Festinger and Maccoby 1964; Freedman and Sears 1965; and Rosenblatt 1966)? Within the Rosenblatt (1966) study, perhaps viewing the dental hygiene slides requires more processing capacity than viewing general psychology slides. A humorous

film could be extensively elaborated and could require more processing capacity than completing a questionnaire (Haaland and Venkatesan 1968). Perhaps manually monitoring a flashing light involves concomitant sub-vocalization and is, in effect, almost identical to vocal-and-manual monitoring (Keating and Brock 1974 and Maccann Haslett 1976).

Even studies that appear at first glance to have utilized the same distraction conditions may not. Baron and Miller (1969) misapply the operationalization of the distractor used by Freedman and Sears (1965), paying attention to the personality of the speaker. In their curious study, the subjects only anticipate hearing a speaker. There can, therefore, be no distraction since there is no message to process nor is there a speaker--consequently, no competition for limited processing capacity. Their study is not a distraction study.

Reported manipulation check measures. Table 2 shows that of 22 studies, 10 utilized no manipulation check. Of the remainder, 3 did not report any values for the measures taken, and one (Keating and Brock, 1974) reported the values for one measure but not for the other.

Even for those studies for which values of the manipulation check are reported along with the significance of the differences between conditions, it is often difficult for the reader to establish whether all of the conditions represent statistically distinct levels of distraction as follow-up tests are not reported. The exceptions are those studies in which there are only two levels of the variable and the Bither (1972) and Bither and Wright (1973) studies in which the distinctiveness of the conditions is reported.

Measures of distraction that have been used. Table 2 shows the measures of the predictor variable that were used in the reviewed distraction studies.

A. Physiological Measures. Heartrate, used by Rule and Rehill (1970) clearly fails as a measure of attention allocated to the distractor. It is not surprising that those subjects who heard an unexpected recording of 22 different sound effects (including a steam-boat whistle, a train, and a race car) experienced an increase in heartrate. Increased heartrate it is probably an indication of arousal, rather than attention allocation.

B. Verbal report measures.

1. Difficulty (Gardner 1966, 1970; and Kiesler and Mathog 1968) is also not a measure of attention devoted to processing a distractor. A person may find it difficult, but possible to avoid allocation of any of the limited processing capacity to the distractor.

2. Effort expended in attending to a message in the presence of a distractor would seem to be an appropriate approximation of the concept of allocation of attention (processing capacity) (Gardner 1966, 1970; Zimbardo et al. 1970; Insko, Turnbull, and Yandell 1974; Petty, Wells and Brock 1976; and Petty 1977). In fact the idea of "effortful" versus "effortless" processing is used synonomously with "processes requiring attention" and "processes not requiring attention" by cognitive psychologists (Posner 1973, p. 40). However, just because researchers consider the terms synonomous, does not mean that research subjects do. They may equate effortfulness with difficulty. Zimbardo et al. (1970) and Insko, Turnbull, and Yandell (1974) used both perceived allocated attention and perceived effort as manipulation checks. Their results indicate that effort is probably seen as being closer to

attention than to difficulty by the subjects. The results of the two measures parallel each other and for both studies, perceived effort is higher for the message-only condition than for the message-set condition (which with the distracting task should have been more difficult). Gardner's (1966, 1970) use of "amount of perceived effort and difficulty" may, therefore, involve two different concepts which may change independently.

The use of: "To what extent did you find the task effortful?" by Petty, Wells, and Brock (1976) and Petty (1977) leaves unspecified just what task is intended and subjects may interpret it as referring to the overall task, to the monitoring task, or to attending to the message. In fact, Petty, Wells, and Brock asked this of non-distracted Ss (1976, experiment 1). Non-distracted subjects reported greater (though perhaps not significantly) effort than those in the low-distraction condition. Notice (Table 2) that as the monitoring task becomes more demanding the "task" is seen as more effortful. This is the opposite of the findings of Zimbardo et al. (1970) and Insko, Turnbull, and Yandell (1974) using "Perceived effort expended in listening to and comprehending the message," and "Effort exerted listening to the communication."

3. Perceived distraction (Osterhouse and Brock 1970; Keating and Brock 1974; Petty, Wells, and Brock 1976; and Petty 1977) suffers from the same shortcomings as difficulty as a measure of allocated processing capacity. A person could consider a distractor very annoying and yet successfully avoid processing it. If perceived distraction is to be used, the form of the question similar to that used in Petty, Wells and Brock (1976) and Petty

(1977): distracted "from paying attention to the message," appears to more closely approach the target concept.

4. Perceived allocation of attention seems to be the best perceptual measure of distraction (Zimbardo et al. 1970; Bither 1972; Bither and Wright 1973; and Insko, Turnbull, and Yandell 1974). Unfortunately, Bither (1972) and Bither and Wright (1973) combine two scales: "degree of attention focused on the message and the amount of the message that was understood." Differences in the amount of the message that was understood does not indicate differential distraction, but a possible effect of distraction. The relationship between these two variables is a topic worth studying, but should not be presumed in establishing whether distraction has been manipulated.

5. Shortcomings of verbal report measures. Regardless of the phrasing of verbal report measures, they all suffer from at least two shortcomings. The first is that they offer the S a chance to second-guess the experimenter in responding. This is probably an even greater problem in experiments where the subjects' following directions dictates a particular response (e.g. Zimbardo et al. 1970). The second problem is even more serious. Ss' subjective experience of the amount of attention allocated to processing a distractor may not correspond to the actual levels allocated.

C. Indirect measures. An intriguing measure of distraction was used by Zimbardo et al. (1970) in an attempt to use "...a distractor which permits a quantitative assessment of the amount of distraction it actually elicits" (p. 671). They measured the change in performance (on the distracting task in the number of problem units from the "speed of

number-summation" task successfully completed) between a warm-up period and the period in which the task and message were competing. The distracting task "...involved totaling the number of letters, rather than the numerical sum, in the names of two digits; such as, 'one plus one equals six.'"

D. A proposed measure of distraction. In the Zimbardo et al. (1970) study, the distractor was self-paced and allocation of attention was voluntary; hence the level of distraction was different for each subject. However, it seems that their approach could be used to independently determine the distraction caused by other distractors. If Ss perform a quantifiable task simultaneously with the distractor, changes in their performance from their performance in a warm-up period from that of a control group could be used as a measure of the distraction caused by the distractor. This would avoid the problems of using self-report, <u>subjective</u> measures of <u>perceived</u> distraction. The search for a standardized task used to quantify different distractors should focus on those that have the following characteristics: a) performance should be readily quantifiable; b) the task should be resource limited, as is elaboration; and c) variance in performance of this task across Ss should be low. This proposed measure of distraction would also avoid any tendency toward consistency in Ss' responses which may result from taking a perceptual measure of distraction from the same Ss who provide the criterion measure. All of the studies reviewed here in which a manipulation check was performed used the same Ss for both measures.

Another possibility would be to use a criterion measure that has been found to be related to distraction in a consistent fashion as a

measure of distraction. However, if this measure were study-specific--as is item-recall--then it could only be used to establish internal validity and not for inter-study comparisons. Another caveat would be that such a measure should always be administered after the same time period following a message, since distraction may interact with time in its effect on recall.

Other Contextual Variables:

There are many contextual variables which have been manipulated besides the presence of distractors.

Forewarning was manipulated by Freedman and Sears (1965), Petty (1977) and Petty and Cacioppo (1977).

Credibility was manipulated by Kiesler and Mathog (1968), Baron and Miller (1969) and Cook (1969). All used verbal descriptions of the sources to effect the manipulation (e.g., parking lot attendant or high-school dropout vs M.D. or Rhodes scholar (Kiesler and Mathog 1968); professor vs retired carpenter (Baron and Miller 1969); and a dental researcher vs a "D" high-school student (Cook 1969).

Kiesler and Mathog (1968) performed a post-test measure of credibility by asking how likeable, fair or biased, and well-informed the source was. They only found a difference on the well-informed measure. However, as they adroitly point out--"As with any after-only design, the check on communicator credibility was confounded with perceptions of the communicator (i.e., derogation) which may have been influenced by the communication" (p. 1128). Of course, the rating could also be influenced by whether or not the Ss had been persuaded.

Baron and Miller also measured source credibility as a multiple post-measure. Their composite index (fair, informed, and intelligent)

showed that there was a significant difference between the two credibility conditions. However, it would be interesting to see if the intelligent measure dominated, since a retired carpenter (low credibility) could be seen as fair or informed, regarding the topic: "...wisdom of special on-the-job training for minority group members" (1969, p. 411).

Cook (1969) also used a post-message measure of the credibility (expertness) of the source. However, the order of structural and credibility measures was counterbalanced. This counterbalancing should remove order effects, but the consistency problems mentioned for the Kiesler and Mathog (1968) study remain a source of confounding. Cook found a significant difference between conditions ($p<0.001$).

Maccoby and Roberts (1972) manipulated opportunity to elaborate. They interspersed 20 second pauses within the message. They also manipulated whether Ss were to write or just think about their responses during presentation. Ss' writing their responses during message presentation could qualify as a distraction manipulation, although they didn't position it as one.

Petty (1977) manipulated body posture (lying, standing or sitting) of the receiver during message presentation.

Cacioppo, Sandman and Walker (1978) manipulated heartrate of the recipient during message presentation.

Wright used message modality (audio vs written) (1973a) and peer-designated, topical opinion leadership (1975) as predictor variables.

Leavitt, Waddell, and Wells (1970) manipulated time-since-eating prior to message presentation.

Brock (1967) manipulated the <u>purpose</u> of the communication by varying <u>persuasive intent</u>. Instructions to Ss in the high-persuasive-intent condition included a statement that "Preliminary copies of the statement were made available on January 10, 1966, with the express purpose of persuading the student body", while those to subjects in the low-persuasive-intent condition were told that the message had been written as a required course assignment (p. 299). Ss responded to a post-measure: "Did you feel the writers were making an active attempt to persuade you?" (p. 301). Chi-square analysis showed that the difference between conditions was <u>non-significant</u> ($p<0.10$). Brock discusses the poor timing of the measure, yet concluded: "Nevertheless, differences in the anticipated direction were obtained and the manipulation was therefore considered adequate" (p. 302). It seems possible that <u>credibility</u> was manipulated instead of <u>persuasive intent</u>, since in one case the message is attributed to a graduate student and in the other to the Faculty Council of the Ohio State University. No credibility measure was taken.

<u>Other Predictor Variables</u>:

Other relationships within the elaboration model besides those involving contextual variables have been examined. Message variables, as well as prior cognitive structure variables have been manipulated so their effect on elaboration and on post-message cognitive structure could be examined.

<u>Message variables</u>:

Cook (1969) and Calder, Insko, and Yandell (1974) varied the <u>number of arguments</u> included in the message. McCullough and Ostrom (1974) varied the <u>number of exposures</u> of similar messages, whereas Cacioppo and Petty (1979) varied the number of exposures to the <u>same</u> messages.

Another message variable that has been manipulated is discrepancy (Brock 1967). However, establishing the discrepancy between the position in the message and that held by the individual requires simultaneous consideration of the prior cognitive structure of the recipient and the position expressed in the message. Discrepancy cannot be established by either one alone. Brock manipulated message extremity and relied on random assignment to conditions to effect the discrepancy conditions that he desired. (See Miller and Levy 1967, footnote p. 160 for a discussion of the issue of extremity vs discrepancy.)

The orientation of the message (counterattitudinal versus proattitudinal) was manipulated by Cacioppo and Petty (1979) and by Petty and Cacioppo (1979, Experiment 1).

Petty, Wells and Brock (1976) and Petty (1977) varied the quality of the arguments used in a message: "One message was designed to contain points that were logically sound, easily defendable, and more compelling than previous tuition increase propaganda...." The other "...was designed to be more open to refutation and skepticism, and be less compelling than previous messages" (Petty, Wells, and Brock 1976, p. 875). They called the variable "message counterarguability" (p. 876). They found a significant difference in the number of counterarguments generated by Ss in the two conditions. However, this was one of their criterion measures as well. Message quality was used again by Petty and Cacioppo (1979), and, again, differences in the criterion measures (attitude and cognitive responses) were cited as evidence of a successful manipulation. Cacioppo, Sandman and Walker (1978) manipulated persuasiveness of the messages which they presented to the subjects. They reported that these were developed in a pretest, but offered no

manipulation check data. It is not clear what the conceptual variable is that has been manipulated in these studies (salience, logic, etc.) and replication by independent researchers would be difficult.

Prior cognitive structure variables:

Several studies have either manipulated or measured variables which involve the cognitive base from which subjects operate in elaborating a communication.

Bither and Wright (1973) measured generalized self-confidence using 15 statements adapted from published personality scales. Wright (1975) used ten items drawn from the same instruments to measure generalized-social-confidence and developed his own ten-item scale to measure information-processing-confidence (p. 5).

Motivational factors (prior cognitive structure variables which could affect the arousal, direction, and persistence of a response (Korman 1974, p. 2)) have been manipulated by several researchers. Maccoby and Roberts (1972) and Roberts and Maccoby (1973) had half of their Ss respond to the criterion measure (level of agreement, on a 15 point scale, with the central argument from the message) prior to the message. However, they variously labeled the variable "motivation to counterargue" (Maccoby and Roberts 1972, p. 264) and "commitment" (Roberts and Maccoby 1973, p. 284). These articles appear to be different views of the same experiment.

Gardner (1966, 1970) manipulated commitment by having Ss choose a movie (for which they would receive tickets) prior to the communication (which was always for the movie which they had rated second).

An evaluative set toward the message was encouraged in certain conditions by Cialdini et al. (1976) and Wright (1973a, 1974b) by

telling Ss that they would be expected to respond to the message content at a future time. Cialdini et al. (1976) told Ss they would be "...involved in a discussion" of the issue (p. 668). Wright told Ss that they would be asked to "...evaluate the product in the advertisement" (1974b, p. 197). His 1973a article is apparently another view of the same experiment.

Osterhouse and Brock (1970) seem to have created a highly reactive situation in their perceived-degree-of-influence manipulation. "Subjects in the high-perceived-influence condition read that the Faculty Committee on Student Tuition was extremely anxious to know the results of this survey in arriving at their final decision" (p. 350). This manipulation should generate an artificially high level of counterarguments that may not be found in passive response situations where, although the issue may be highly involving, individuals are not in a position to influence someone else with their responses. Three separate measures of whether a) student opinion, b) opinions from the survey, and c) the S's opinions would affect the decision of the Faculty Committee all revealed significant differences between conditions.

In the same experiment (Osterhouse and Brock 1970), communication threat was also manipulated by "...varying the likelihood that the unwanted event [tuition increase] would occur" (p. 350). However, in doing so, these researchers may have inadvertently manipulated the credibility of the source: "...subjects in the strong-threat condition then read that the committee had been favorably impressed by the report," whereas "subjects in the mild-threat condition were informed that, while the committee was deeply appreciative of the work done by Professor Patton [the report's purported author], most of the committee

members had reservations about the desirability of a tuition increase at the present time" (p. 350). No measure of credibility was taken. Ss in the high-threat condition were significantly more _concerned_, more _uncomfortable_, more _threatened_, and saw the tuition increase as more _likely_ than those in the mild-threat condition.

Petty and Cacioppo (1979) manipulated _issue involvement_ (as opposed to _response_ involvement) by varying whether the advocacies [regarding mixed-sex visitation hours (Experiment 1) or required comprehensive exams for all seniors (Experiment 2)] applied to the subjects' or to another university. In Experiment 1 the authors utilized a manipulation check question (asked after the criterion measures) so that subjects could indicate how "involving" the communication was on an 11-point scale. Subjects in the high involvement condition rated the communication as significantly more involving than did those in the low involvement condition ($p < 0.03$). The respective means were on opposite sides of the scale midpoint ($M = 6.67$ and $M = 4.75$). However, in Experiment 2 Petty and Cacioppo (1979) utilized a different manipulation check measure for involvement. They asked subjects (after all but the recall criterion measures) to respond to an 11-point scale indicating "...how much thought they put into evaluating what the speaker had to say" (p. 1921). However, they defined issue involvement as "...the extent to which the attitudinal _issue_ under consideration is of personal importance..." (p. 1915). While the personal importance of an attitudinal issue may sometimes be related to how much thought is engaged in to evaluate what a speaker has to say, they do not provide evidence that this relationship is reliable enough to use this manipulation check measure as a measure of issue involvement.

Rule and Rehill (1970) manipulated self-esteem by giving varying feedback regarding performance on a maze-solving task (p. 361). "Heart-rate and other data" (footnote, p. 361) from a previous study indicated differences between conditions as did heart-rate data in this study. What the heart-rate measure reflects is unclear.

A final manipulation of prior cognitive structure (Brock 1967) was the presence or absence of a "priming" counterargument printed as an "example" on the page on which Ss were to record their thoughts.

Criterion Measures

As criterion measures of the effect of message, context, and prior-cognitive-structure variables, processing (elaboration) and structural (cognitive structure) variables have been used, depending on the relationship of interest. In some cases, processing variables have also been considered the predictor variables in studies examining their relationship with structural variables. Table 3 presents different aspects of both processing and structural measures that have been used. The significance of the dimensions which are presented will be explained in the following discussion.

Processing Variables as Criterion Measures:

Mode of collection:

Measuring elaborative responses has essentially consisted of the collection of verbal protocols. The use of psychophysical measures has not been exploited. One reason is that the measures are ambiguous. The researcher cannot tell what sub-vocalizations may be indicated by such measures (e.g., "...counterargumentation as opposed to some other language-related process (e.g., silently repeating the message).

Table 3. Dependent Variable Timing, Time Limits and Collection Mode.

Authors and Year	Processing Measure Label	Processing Measure Timing	Time Limit	Collection Mode	Structural Measure Label	Structural Measure Timing
Festinger and Maccoby 1964	none	---	---	---	attitude	immediately following message
Freedman and Sears 1965	none	---	---	---	attitude	immediately following message
Rosenblatt 1966	none	---	---	---	1. attitude 2. recall	2nd postmeasure 4th postmeasure
Gardner 1966	none	---	---	---	1. desirability 2. attitude 3. recall	1st postmeasure 2nd postmeasure 3rd postmeasure
Krugman 1966-67	connections	followed questions about surrounding editorial material	?	oral	none	---
Brock 1967	precommunication thoughts	prior to message	10 min.	written	1. agreement 2. convincingness 3. recall	1st postmeasure 2nd postmeasure 5th postmeasure
Vohs and Garrett 1968	none	---	---	---	attitude	immediately following message
Kiesler and Mathog 1968	none	---	---	---	communication effectiveness (sum of agreement and convincingness of speaker)	5th and 6th items on page following page for recall (recall data not reported)
Haaland and Venkatesan 1968	none	---	---	---	1. attitude 2. recall	1st postmeasure 2nd postmeasure
Venkatesan and Haaland 1968	none	---	---	---	recall	immediately following message
Silverman and Regula 1968	none	---	---	---	persuasibility (5 items combined)	immediately following message

(continued)

Table 3 continued.

Authors and Year	Processing Measure Label	Processing Measure Timing	Time Limit	Collection Mode	Structural Measure Label	Structural Measure Timing
Greenwald 1968 a. Cullen	cognitive responses	immediately following message	?	written	opinion (4 items combined)	followed thought listing
b. Love	1. cognitive reactions 2. opinion (thought listing) 3. opinion (thought listing)	interposed (after each of 3 paragraphs) followed last-paragraph response one-week delay (more likely a structural measure - therefore also listed in that column	? ? ?	written written written	1. retention of communication content 2. retention of cognitive reactions 3. opinion (thought listing)	1. after thought listing (immediately and one-week delay) 2. immediately after content recall (immediately and one-week delay) 3. one-week delay
Baron and Miller 1969	counterargument	followed credibility, attitude and agreement measures (prior to message)	1 min.	written	attitude (sum of 4 items)	2nd postmeasure
Cook 1969 Experiment 1	comments	? opportunity concurrent with reading	1 min.	written	1. attitude (3 separate items) 2. recall	varied varied
Experiment 2	same as Exp. 1	same as Exp. 1 plus a request for additional comments at the end of the message	?	written	1. agreement 2. recall	1st post-comment measure 4th post-comment measure
Zimbardo et al. 1970 Experiment 1	none	---	---	---	1. learning (recall) 2. attitude	1st post-message measure 3rd post-message measure
Experiment 2	"counterarguing" - "relative percentage of the time given to listening and to arguing"	?	N.A.	scale	1. recall 2. attitude (agreement) 3. attitude (evaluation of benefit of the proposal)	1. 1st post-message measure 2. sometime after the recall measure 3. sometime after the recall measure

(continued)

Table 3 - continued.

Authors and Year	Processing Measure Label	Processing Measure Timing	Time Limit	Collection Mode	Structural Measure Label	Structural Measure Timing
Zimbardo et al. 1970 Experiment 3	same as Exp. 2	?	N.A.	scale	same as Exp. 2	same as Exp. 2
Osterhouse and Brock 1970 Experiment 1	counterarguments	followed attitude measure	3 min.	written	1. acceptance (ave. of 3 items) 2. recall	1. 1st post-meassage measure 2. followed thought-listing
Experiment 2	counterarguments	followed acceptance measure	3 min.	written	1. acceptance (ave. of 2 items) 2. recall	1. 1st post-message measure 2. on 2nd questionnaire followed attitude and thought-listing)
Rule and Rehill 1970	none	---	---	---	1. agreement 2. retention (responses to specific questions)	1. immediately after message 2. after agreement measure (these measures followed recall of sound effects for half of subjects)
Leavitt, Waddell and Wells 1970	Personal Product Response	? (after commercial)	?	oral	? (possible recall measure to verify reception of commercial	?
Gardner 1970	none	---	---	---	1. desirability 2. attitude (sum of 5 items) 3. recall	1. 1st post-message measure 2. 2nd post-message measure 3. 3rd post-message measure
Maccoby and Roberts 1972	counterarguments	1. during message (message with or without 20-sec pauses) 2. post-meassage, following attitude measure	?(20-sec pauses ?	written written	attitude (agreement)	prior to message for "committed" subjects; post-message for all subjects

(continued)

Table 3 - continued.

Authors and Year	Processing Measure Label	Processing Measure Timing	Time Limit	Collection Mode	Structural Measure Label	Structural Measure Timing
Keating and Latane 1972	none	---	---	---	attitude (sum of 17 items)	immediately after message
Bither 1972	none	---	---	---	1. attitude 2. learning (recall)	1. 1st post-message measure 2. 3rd post-message measure
Wright 1973a, 1974b, 1975	thoughts	immediately after	3 min.	written	1. convincingness 2. attitude 3. buying intentions 4. recall	1. 1st post-listing measure 2. 2nd post-listing measure 3. 3rd post-listing measure 4. 4th and 5th post-listing measure
Roberts and Maccoby 1973	essentially a re-working of the data from Maccoby and Roberts (1972)					
Bither and Wright 1973	none	---	---	---	1. attitude (sum of 4 dimensions) 2. recall	1. immediately after message 2. 2nd post-message measure
McCullough and Ostrom 1974	thought, comment or opinion	during and after presentation	none	written	none	
Keating and Brock 1974	thoughts and ideas	3rd post-meassage measure	3 min.	written	1. agreement 2. appropriate tuition 3. recall	1. 1st post-message measure 2. 2nd post-message measure 3. followed thought-listing
Insko, Turnbull and Yandell 1974	thoughts	alternated - before or after attitude measure (immediately after message	?	written	1. attitude a. agreement b. sum of 3 items 2. recall	1. these two measures alternated before or after thought-listing 2. 6th post-message measure
Calder, Insko and Yandell 1974 (Experiment 3)	thoughts	immediately after message or with 1 or 2 week delay (prior to belief for half of the subjects)	10 min.	written	belief (certainty of innocence or guilt)	altenately before or after thought-listing

(continued)

Table 3 - continued

Authors and Year	Processing Measure Label	Processing Measure Timing	Time Limit	Collection Mode	Structural Measure Label	Structural Measure Timing
Petty, Wells and Brock 1976 Experiment 1	cognitive responses	alternately before or after the attitude measure (immediately after message)	2½ min.	written	1. agreement 2. appropriate tuition 3. recall	1 & 2 either preceded or followed thought-listing 3. followed thought-listing and 1 & 2 above
Experiment 2	cognitive responses	followed 2 attitude measures	2½ min.	written	1. agreement 2. appropriate tuition 3. recall	1 & 2 followed the message and preceded thought-listing 3. 8th post-message measure
Maccann Haslett 1976	thoughts and ideas	followed 6 attitude measures	?	written	1. attitude (composed of 6 agreement measures) 2. comprehension (3 items)	1. immediately after message 2. followed attitude measure
Cialdini et al. 1976 (Experiment 1)	thoughts	pre-discussion (after subjects had stated their position and involvement on two issues. Elicited 1st for Issue 1	3 min.	written	1. position (for or against) 2. opinion (4 items) 3. opinion (8 items)	1. 1st pre-message measure 2. followed thought-listing 3. followed 2 above
Petty 1977 Experiment 1	thoughts	1st measure (pre-message	3 min.	written	agreement	immediately after message
Experiment 2	thoughts	after 2 structural measures	2½ min.	written	1. agreement 2. appropriate tuition	1. 1st post-message measure 2. 2nd post-message measure
Experiment 3	thoughts	after agreement measure	2½ min.	written	agreement	1st post-message measure

(continued)

Table 3 - continued.

Authors and Year	Processing Measure Label	Processing Measure Timing	Time Limit	Collection Mode	Structural Measure Label	Structural Measure Timing
Petty and Cacioppo 1977						
Experiment 1	ideas	1st measure (pre-message)	2.5 min.	written	agreement	1st post-message measure
Experiment 2	thoughts	1st measure (pre-message)	3 min.	written	agreement	1st post-message
Cacioppo, Sandman, and Walker 1978	thoughts	followed multiple messages which had accompanying structural measures	?	written	agreement	followed each message
Cacioppo and Petty 1979	cognitive responses	followed agreement measure	3 min.	written	1. agreement 2. recall	1. 1st post-message measure 2. followed thought-listing
Petty and Cacioppo 1979						
Experiment 1	cognitive responses	followed structural measures	2.5 min.	written	1. four semantic differential scales 2. agreement	1. 1st post-message measure 2. followed semantic differential items
Experiment 2	cognitive responses	followed semantic differential and agreement measures	2.5 min	written	1. four semantic differential scales 2. agreement 3. recall	1. 1st post-message measure 2. 2nd post-measage measure 3. 6th post-message measure

However, using subvocal measures in conjunction with the open-ended measures discussed at length above might well circumvent this problem and, therfore, would seem to be a logical next step in counterargumentation research" (Miller and Baron 1973, p. 115). Cacioppo, Sandman and Walker (1978) monitored chim EMG activity during message presentation. Although they also collected verbal protocols of reported elaboration, they did not report any examination of a relationship between the physiological measure and the verbal measure.

The two <u>self-generated</u> means of collecting verbal protocols are having the Ss either speak or write their thoughts. Speaking has the advantage of spontaneity and freedom from the physical activity and apprehension (spelling, grammar, etc.) involved in writing. "Vocal expression is a more fluid and direct means for translating from subvocal thought; we think verbally rather than in print" (Wright 1974a, p. 235). However, as Wright also points out, "stage-fright" may accompany oral recitation of thoughts (1973. p. 56). It seems that recitation could also <u>encourage</u> those who <u>enjoy</u> the theatrical aspects of the proceedure. A later section will specifically treat the reactivity of the protocol elicitation process.

A review of Table 3 reveals that almost all protocol-elicitation studies have utilized <u>written</u> responses. Maccoby and Roberts (1972) state that this procedure "...seemed to be the best available method for getting close to what goes on in a receiver's head during reception. While the writing process is obviously at least a step removed from the thought process, per se, given the present state of technology it is the best technique we have" (p. 263).

From a purely practical, experimental point of view, writing also eliminates the need to have the responses transcribed and makes possible group administration of experimental treatments.

A final means of collecting elaborative responses that has been suggested uses responses that are externally generated. Miller and Baron (1973) suggested the use of recognition of counterarguments as a measure of elaboration: a subject "...would be given a list of counter-arguments...[and]...would be asked to check off within the time limit imposed, those counterarguments that occurred to him during the presentation of the communication" (p. 113). However, this approach seems highly reactive as Ss may just check arguments that seem reasonable and appropriate as they read them, whether or not the arguments had spontaneously occurred to them.

Timimg of the elaboration measure:

Protocols may obviously be collected prior to, during, or after message presentation. Those collected after message presentation may be collected immediately, after an intervening measure (during the same session), or at a later date. Table 3 shows that most protocols have been collected after message presentation. There are operational and conceptual issues surrounding the timing of protocol collection.

Protocol measures taken prior to message exposure. Anticipatory elaboration was studied by Brock (1967); Baron and Miller (1969); Cialdini et al. (1976); Petty (1977); and Petty and Cacioppo (1977). These pre-message protocols must be measuring something different from what is intended in many of these studies, i.e., cognitive responses that occur in response to the communication. "Although receivers' pre-message comments can be defined as arguments for or against a

particular issue, it is difficult to conceive of them as counter-arguments in the strict sense of the term" (Roberts and Maccoby 1973, p. 280). Their occurrence also "...does not indicate whether or how a receiver might actually use such counters" (Maccoby and Roberts 1972, p. 263). Anticipatory responses probably reveal the cognitive structure that the recipient brings to the communication situation, rather than exposing the response process that is hypothesized in the elaboration model.

Protocol measures taken during message exposure. Conceptually, eliciting verbal protocols during exposure is most compatible with the elaboration model. This procedure attempts to capture the responses as they occur, revealing the elaboration process as the person responds to the message. In this way, the elicitation procedure can avoid "...problems associated with recall in post-message interviewing" (Miller and Baron 1973, p. 112). Measuring elaborative responses during exposure may be especially appropriate for long messages as "...an after-exposure recording may take the subject's memory of his own responses or favor responses occurring near the end of the message" (Wright 1974a, p. 232).

In addition to the problem of recall, measuring elaboration after exposure "...may also simply be telling us that this is how the receiver rationalizes his position after the fact. In short, there seems to be no substitute for obtaining counters being used at the time of message reception" (Maccoby and Roberts 1972, p. 263). The major disadvantages of eliciting protocols during exposure is the reactive and disruptive nature of the procedure. "Clearly, any elicitation technique that forewarns the subject prior to message exposure that a thought recording

measure is imminent or that requires the subject to interupt message transmission to record thoughts partially compromises the validity of the measure" (Wright 1974a, p. 229-230).

If protocols are elicited during exposure, they should be produced vocally so that the additional task of writing will not occupy crucial portions of the limited processing capacity.

Love, as reported in Greenwald (1968), Cook (1969), Maccoby and Roberts (1972), and McCullough and Ostrom (1974) elicited protocols during or intermittently with message presentation (see Table 3).

Protocol measures taken immediately following message presentation. Measures of elaboration that follow the message in the same experimental session may occur immediately after message presentation or may follow other activities (attitude, recall measures, etc.). Eliciting protocols immediately after the message minimizes problems of recall and consistency with other post-message measures. To minimize reactivity and a resulting artificial inflation in thoughts listed, Wright (1974a, p. 232) suggested that the administrations of a post-measure be administered unexpectedly. Of course, for unexpected measures, the instructions must intervene between the message and the measure of elaboration.

Cullen, as reported in Greenwald (1968); Love as reported in Greenwald (1968); Wright (1973a, 1974b, 1975); McCullough and Ostrom (1974); Insko, Turnbull, and Yandell (1974); Calder, Insko, and Yandell (1974); and Petty, Wells, and Brock (experiment 1, 1976) utilized elaboration measures that immediately followed message presentation.

Protocol measures taken following other post-message measures. If other measures precede post-message protocol collection, there is the

possibility that responses will be generated to maintain consistency with those earlier measures. However, as Miller and Baron (1973) point out: "It is always possible that regardless of how the researcher temporally orders his measures of counterarguing and attitude change, people construct counterarguments to support whatever position they hold <u>after</u> they have been exposed to persuasive materials" (p. 107).

Because of the interference of other measures with the recall of elaborative responses that may have occurred during exposure, it is unlikely that protocols that are elicited after intervening measures disclose <u>actual processing responses</u> that occurred during exposure. "If other activities occur after a message and before thought listing, thought listing proceedures must necessarily tap long-term memory" (Calder 1978, p. 632). In the studies by Baron and Miller (1969); Osterhouse and Brock (1970); Maccoby and Roberts (1972); Keating and Brock (1974); Insko, Turnbull, and Yandell (1974); Petty, Wells, and Brock (1976); Maccann Haslett (1976); Petty (1977); Cacioppo, Sandman and Walker (1978); Cacioppo and Petty (1979); and Petty and Cacioppo (1979) other measures intervened between the message and protocol elicitation.

<u>Protocol measures taken in later sessions</u>. <u>Delayed</u> measures of elaboration are even more suspect than measures made after an intervening structural measure as an indicator of thoughts that may have occurred <u>during</u> message exposure. They are probably indicators of the resultant, underlying propositional structure which is comprised of both the interpretation of the message and responses to the message. Love (Greenwald 1968); Calder, Insko, and Yandell (1974); and Wright (1974b) reported delayed elicitation of elaboration protocols.

Time allowed for elaboration listing:

In their desire to have only those thoughts listed which had occurred during exposure to the communication and avoid the intrusion of thoughts generated during the listing task, some researchers have limited the time which Ss have available for the listing task. (See Wright 1974a, p. 234). "Unfortunately...a short time period does not eliminate the possibility that subjects respond to the measure by producing counterarguments that are readily available to them but not spontaneously thought of during or before the message" (Miller and Baron 1973, p. 106).

In natural settings, message recipients have an indefinite amount of time in which to elaborate after message presentation. To increase the generalizability of research findings, Miller and Baron (1973) suggested giving "...the subject an analogous amount of time in the setting [following message presentation] and then requiring him to produce his counterarguments in a short period of time" (p. 107). However, it is unclear what an "analogous amount of time" might be.

Table 3 shows that many different response times have been utilized in the studies cited. Wright (1974a, p. 234) suggested that a standardized time for uniform message presentations (such as broadcast commercials) would facilitate inter-study comparisons.

Protocol coding:

Selection of relevant dimensions for protocol coding. "One of the beauties of the thought-monitoring approach is its indirect manner of extracting the response variables of interest. [Subjects]...report all thoughts...in ignorance of what type of response the researcher will be singling out for examination" (Wright 1974a, p. 235). Even though the

thought-monitoring approach may be _unbiased in its selectivity_ (the effect of instructions to the subjects will be discussed in a later section), researchers have not been. The central focus of most studies has been on _counterarguing_ as an elaborative response. "The prominance of Festinger and Maccoby's (1964) counterargument disruption hypothesis stems in part from its primacy, but it also reflects the importance of the counterargument concept in explaining numerous attitude change phenomena" (Baron, Baron, and Miller 1973, p. 316). "...Regardless of whether our comments are overt or covert, regardless of whether the persuader is present or absent, regardless of whether or not counter-arguments facilitate resistance to the appeal, it seems that one characteristic way of dealing with a persuasive message is to argue with it" (Roberts and Maccoby 1973, p. 260).

However, counterargument has certainly not been the only category used for coding responses, nor is it the only conceptually important category in the persuasion process. "...We would contend that the configuration of all comments, counter or supportive which a receiver thinks up and/or rehearses during reception will exert strong influence on the final response" (Maccoby and Roberts 1972, p. 262).

Calder (1978) suggests the inclusion of imagery as a legitimate response mode that might be coded. As a way of getting closer to imagery that may be evoked in responding to a message, he suggests considering cognitive responses as "scripts"--a term used by Abelson (1976) (see also Schank and Abelson 1975). "By 'script' I mean a _coherent sequence of events expected by the individual involving him either as a participant or as an observer_. ...The basic ingredient of scripts I label a _vignette_. This is an encoding of an event of short

duration, in general including an image (often visual) of the perceived event and a conceptual representation of the event. ...The perceptual image might include codes from various sense modalities, and also codes for experienced affect" (Abelson 1976, pp. 3334). This concept is essentially comparable to the schema that comprise episodic memory as discussed in Chapter 2, except that Abelson notes that many <u>scripts</u> are relatively standardized and uniformly recognized within a culture.

Viewing cognitive responses as scripts will require new coding schemes. "...The notion of scripts also calls attention to the need to view cognitive responses as ordered, contextual, and autobiographical" (Calder 1978, p. 633).

<u>Protocol coding categories</u>. The categories into which protocol statements have been coded (Table 4) involve several dimensions. The most common has been <u>directionality</u> or <u>favorability</u>.

Directionality has been utilized within another dimension: the <u>target</u> of the statement (i.e., the <u>message</u>; an <u>argument</u> or <u>point</u> expressed in the message; the <u>source</u> of the message; and in the case of "personal-product-response" (Leavitt, Waddell, and Wells 1970), the <u>product</u> in an advertisement).

To reflect the fact that some statements are more positive or more negative than others, <u>intensity</u> is another dimension that has been used.

The <u>source</u> of the statement is a dimension that reflects whether the statement apparently came directly from the message, represents a modification of part of the message by the recipient, or was purely recipient generated. This a significant dimension vis-a-vis the elaboration model. Statements which are directly from the message are less likely to have been incorporated into the cognitive structure than are recipient-modified or recipient-generated statements.

Table 4. Protocol Coding Approaches (N = number, P = proportion)

Authors and Year	Subject or Judge	Categories	Indices
Krugman 1966-67	both	1. connection (positive and negative (judge) 2. time (during ad vs during (subject)	1. N of connections/second 2. average N of connections
Brock 1967	judge	counterarguments	N of counterarguments
Greenwald 1968 A. Cullen	both	1. favorability, 2. degree, 3. source (externally generated. recipient modified, recipient generated)	1. N in each source category 2. directional content index (Σ weights for advocated position minus Σ weights for opposing position divided by Σ all weights)
B. Love			
1. "cognitive reactions"	?	favorable and unfavorable	favorable minus unfavorable
2. opinion (thought-listing)	?	1. directionality	?
		2. source (message, "cognitive reactions", neither)	?
Baron and Miller 1969	judge	1. N of words 2. N of complete thoughts	1. N of words 2. N thoughts

(continued)

Table 4 (continued)

Authors and Year	Subject or Judge	Categories	Indices
Cook 1969	judge	1. agreements 2. counterarguments	1. P of subjects agreeing with at least one argument 2. N of counterarguments
Osterhouse and Brock 1970	judge	counterarugments	N of counterarguments
Leavitt, Waddell and Wells 1970	judge	personal-product-response	presence or absence of personal product response
Maccoby and Roberts 1972	judge	I. supportive or counter: A. specific to: conclusion, source/communication, point made in the message B. subject generated arguments C. comments about the issue (all statements in category I were rated for intensity). II. irrelevant	1. total N of comments 2. N of support comments 3. N of counter comments 4. algebraic sum of intensity scores

(continued)

Table 4 (continued)

Authors and Year	Subject or Judge	Categories	Indices
Wright 1973a, 1974b, 1975	both	1. counterargument, source support argument, curiosity (judge) 2. importance of each thought (subject) 3. source: advertising-originated, recipient-modified, recipient-generated (subject)	1. N of counterarguments, support arguments, source derogations 2. weighted (by importance) counterarguments, support arguments and source derogations
Roberts and Maccoby 1973	judge	same categories as Maccoby and Roberts (1972)	1. P total comments 2. P supportive comments 3. P counter comments 4. N total comments 5. N supportive comments 6. N counter comments 7. N supportive minus counter 8. algebraic sum of intensity scores 9. algebraic sum of intensity scores divided by N comments
McCullough and Ostrom 1974	judge	favorable, unfavorable, neutral	N positive minus N negative
Keating and Brock 1974	judge	counterarguments	N of counterarguments

(continued)

Table 4 (continued)

Authors and Year	Subject or Judge	Categories	Indices
Insko, Turnbull and Yandell 1974	both	1. favorability (subject) 2. source: externally-oriented, recipient modified, recipient-generated (subject) 3. counterarguments (judge)	1. N favorable minus N unfavorable within each source category 2. N of counterarguments
Calder, Insko and Yandell	subject	1. favorability: prosecution, defense, neutral 2. degree of favorability 3. origin: external, recipient modified, recipient generated	1. N of prosecution thoughts 2. N of defense thoughts 3. prosecution plus defense 4. prosecution minus defense 5. N in each source category 6. sum of weights for each source category
Petty, Wells and Brock 1976	both	favorable, opposed, neutral (only subject codings analyzed)	1. N favorable thoughts 2. N opposition thoughts
Maccan Haslett 1976	judge	counterarguments	N of counterarguments
Cialdini et al. 1976	subject	1. favorable, opposed, neutral 2. confidence in the validity each thought	N of favorable thoughts

(continued)

Table 4 (continued)

Authors and Year	Subject or Judge	Categories	Indices
Petty 1977			
Experiment 1	judge	counterargument, proargument, neutral	N in each category
Experiment 2	judge	same as Experiment 1	1. N of counterarguments 2. N of proarguments
Experiment 3	judge	same as Experiment 1	same as Experiment 2
Petty and Cacioppo 1977			
Experiment 1	both	counterargument, favorable thought, neutral (only judge's coding used in reported analysis, except when judges disagreed)	N in each category
Experiment 2	both	same categories as Experiment 1 (only judge's codings reported)	N in each category
Cacioppo, Sandman and Walker 1978	judge	counterarguments, favorable thoughts, neutral thoughts	N in each category
Cacioppo and Petty 1979	judge	counterarguments, favorable thoughts, neutral/irrelevant	N in each category

(continued)

Table 4 (continued)

Authors and Year	Subject or Judge	Categories	Indices
Petty and Cacioppo 1979			
Experiment 1	subject	in favor, opposed, neutral/ irrelevant	N in each category
Experiment 2	subject	same as Experiment 1	N in each category

Two other dimensions have been restricted to subject-coding: the importance of the statement to the S and the confidence that the S has in the statement.

A study reported by Brock has had a tremendous influence on research into elaborative responses. He developed and used an operational definition for counterargument:

> For an idea to be scored as a counterargument it had to be a declarative statement directed specifically against the advocated tuition increase. Furthermore it had to mention a specific unfavorable or undesirable consequence that was not simply a restatement or paraphrase of the fact of a tuition increase... Not counted as counterarguments were simple statements of opposition, ...affective reaction,...all interrogative statements, ...attack on a target other than the tuition increase, ...or paraphrase with or without simple opposition... Of course, positive statements, ...and alternative proposals and accomodations... were never counted as counterarguments (1967, p. 301).

His operational definition was used intact by Osterhouse and Brock (1970, Experiment 1) and by Insko, Turnbull, and Yandell (1974).

Brock's original operational definition was modified by Osterhouse and Brock (1970, Experiment 2) to make it more inclusive. "...The scoring of counterarguments was extended to two other kinds of statements: statements which suggested alternative methods of raising money to finance needed improvements...and statements which challenged the accuracy or validity of the communication" (Osterhouse and Brock 1970, p. 350). This extended version was used by Keating and Brock (1974) and Maccann Haslett (1976), and according to Wright (1974a), served as the basis for the counterargument operational definition that he developed and used for his study reported in Wright (1973a, 1974b, 1975; see 1973a, Appendix A, p. 62 for his operational definition).

Petty, Wells, and Brock (1976) modified the Osterhouse and Brock (1970) version to include statements of affect which were specifically

excluded in the 1970 version. This modification was used by Cacioppo, Sandman and Walker (1978) and by Cacioppo and Petty (1979).

Practical problems in protocol coding. The coding of verbal protocols requires several decisions. Should it be done by judges or by the subjects? What categories should be used? Should the categories be selected *a priori* on a theoretical basis or should they be determined on the basis of content analysis of the responses? What is the appropriate quantitative representation of the nature of the response (e.g., total number of thoughts, number of thoughts within one category, number of positive minus the number of negative thoughts, ratio of one type to some other type, etc.)?

A. Subject vs. judge coding of protocols. Subject-coding has both conceptual and practical appeal. Ss should know if they intended a protocol statement as an expression of a positive or negative viewpoint. Also, once the protocols are collected, they are already coded. This could be an important consideration for researchers who want to perform the analysis quickly or who don't have the personnel available to perform the coding. Judge coding may be very time-consuming. "Here we should point out a disadvantage of this type of research: it took three full-time assistants over two months to develop the coding scheme and to code all of the protocols" (Maccoby and Roberts 1972, p. 269).

A disadvantage of subject-coding is that Ss may not appreciate subtle, but theoretically important distinctions in operational definitions for categories and they may not assign their statements to the same category as the researchers would. If judges are used, inter-rater reliability can be assessed and multiple judgements may be used for category assignment. (See Wright 1974a, p. 235, for a related

discussion). Table 4 indicates which studies used subject-coding and which used judges.

B. Theoretical categories vs content analysis for protocol coding. All of the studies reported here used theoretically-determined categories for protocol coding, except those of Maccoby and Roberts (1972) and Roberts and Maccoby (1973). The final categories for those two reports were developed by examining protocols from pilot studies and from their experiment. "After a number of weeks of hard work and 'discussions' conducted at rather high decibel levels, we finally settled on six content categories which seemed to cover all of the data, to have a good deal of face validity, and to produce acceptable inter-coder reliability" (Maccoby and Roberts 1972, p. 267).

The primary advantage of such content-based coding schemes is that they can uniquely characterize the set of responses from which they were developed. When different studies yield idiosyncratic response patterns, this approach can avoid forcing the responses into preconceived categories. However, the uniqueness of the approach is also its primary disadvantage. Cross-study comparisons are not possible <u>within</u> categories when this approach is used. Of course, the <u>time</u> required for the development of the coding scheme is another drawback.

When consistent, theory-based categories are used for coding protocols, hypotheses derived from the theory can be tested and the findings from different studies may be compared. However, the consistency may be illusory, since operational definitions, having the same label, which are used in a series of studies may include subtle evolutionary changes as has been exhibited within the Brock tradition.

Another disadvantage to the consistent use of theory-based categories is that it may unnecessarily restrict the responses that are

seen as significant. (See Ward 1974 for a related discussion). Table 4 presents the coding categories used in the studies reviewed here.

<u>The quantitative representation of response protocols</u>. Once protocols have been coded, there are many possible ways of using the counts within the various categories in order to characterize the responses.

One possibility is to simply analyze the <u>number</u> of responses within categories, treating them singly, in a univariate fashion, or using some multivariate proceedure to consider several categories simultaneously.

Another possibility is to arrive at some composite representation, by subtracting the number in one category from the number in another, or by deriving some ratio (dividing by either the total number or the number in one category). Extreme caution should be observed in deriving composite variables since the result should have <u>conceptual</u> as well as mathematical meaning. Table 4 illustrates the various indices utilized in the studies which are reviewed here.

These indices vary in their appropriateness for representing the nature of Ss' responses. The appropriateness of the index depends upon the relationship being studied, as this will dictate what aspect of the response pattern is relevant. Appropriateness of an index also depends upon practical concerns, such as the error variance which may be expected in the data across all respondents. For example, dividing by the total number of responses for an individual will reduce the error variance introducted by verbose or reticent respondents.

Indices which represent the <u>total number of words</u> or <u>thoughts</u> would seem to be a measure of activity or of arousal. This could be useful if the effect of distraction or of motivating aspects of the situation or of the message were being studied. However, the number of words or

thoughts is not an adequate measure of _how_ the message was incorporated into the recipient's cognitive structure. A simple word or thought count cannot represent the evaluative content of Ss' responses.

Indices which represent the _number of positive or of negative thoughts or ideas_ expressed in a protocol provide additional insight into the evaluative nature of the Ss' responses. However, since these indices only report positive _or_ negative thoughts, they fail to show the overall (possibly balanced) nature of the response. However, they could be useful if the target criterion in a study was just negative or just positive responses.

Indices which involve an _algebraic sum_ of positive and negative responses can represent the overall positive or negative nature of the response. However, this assumes that the negative and positive responses are equally salient to the subject and that positive and negative thoughts compensate for each other. Such a representation would presume that if the number of negative thoughts in a S's protocol exceeded the positive by one, then that S must have successfully resisted the message. It may be that one, especially salient, positive thought could compensate for many negative ones. This could be especially likely if the Ss felt compelled to give a balanced treatment to both negative and positive points in the protocol to show the experimenter how rational and open-minded he/she was.

Indices which involve simple counts or algebraic sums suffer from the additional shortcoming of giving inflated values of the index for loquacious respondents (e.g., if an algebraic sum were used, a S with 6 positive and 4 negative statements would be assigned a +2, whereas an S with 3 positive and 2 negative would be assigned a +1, yet both Ss would have the same proportion of positive to negative statements.

The use of counts or sums also inflates the error variance in the responses across all Ss and thereby reduces the sensitivity of statistical tests that may be used in the analysis. Yet the variance in wordiness does not provide information relevant to the evaluative nature of the response.

Weighting schemes used in conjunction with algebraic summation can provide Ss the opportunity to rate the relative importance of their thoughts to them and may result in more realistic representations of responses. Thoughts that had been added just to give a rationally appearing response may be (but not necessarily will be) given lower weights of importance. However, even though Ss may perform the weighting task, they may not know which thoughts had the greatest relevance to them. Again, they may rate those thoughts highly which they think the experimenter would consider important.

Another possible weighting scheme is to have judges apply some operational criteria in assigning weights to the various thoughts (e.g., the intensity ratings used by Maccoby and Roberts 1972 and by Roberts and Maccoby 1973). This has the advantage of consistency in applying a uniform scheme across all Ss' responses. Of course, judges cannot know how strongly any one S felt in expressing a thought.

Ratio indices can represent the relative number of thoughts which were positive, negative, or neutral. There are choices as to what should be represented in the numerator: number of positive or negative responses, or a simple or a weighted algebraic sum. If the experiment is testing the effects on a particular type of response (i.e., positive or negative), then that response type would be the natural choice for the numerator. If the study is examining the effect on the overall

nature of the response, then an algebraic sum may be appropriate. This would be intuitively appealing as a balanced response would be assigned a 0 (neutrality). But, again, an algebraic sum may be misleading since all statements may not be equally salient to the respondent. Weighted sums may be used to provide more information.

The researcher must also decide which variable(s) should appear in the denominator. If the numerator were the number of positive statements, then the denominator could be the number of negative statements, or vice versa. More information would be provided if the total number of responses were included in the denominator (including neutral statements). Notice in Table 4 that Cullen (reported in Greenwald 1968) used the sum of the weights assigned to the statements as the denominator. While this has the advantage of putting the numerator and denominator in the same units, it cannot include information regarding neutral comments.

Protocol categories and expectancy-value formulations. Several researchers have used, or suggested, coding and combining protocol data in a way that is compatable with expectancy-value formulations. Wright (1973a) used "weighted integrative models" (p. 58) which involved summing the product of weights of importance assigned by the Ss and thoughts expressed within a category. In this case, the numerical value of each thought would be one.

Calder, Insko, and Yandell (1974) likened their approach to Fishbein's multiattribute model: "Likewise, a specific belief (B_i) might be represented by a thought S wrote down where the strengths of the beliefs were assumed to be the same (+1). The affect (a_i) component is roughtly analogous to whether a thought was listed as favoring the

prosecution (+1) or the defense (-1)...As in Fishbein's model the two components, in effect, were multiplied and summed over all thoughts. The result was an index that was a S's predicted overall belief as derived from his thoughts" (Calder, Insko, and Yandell 1974, p. 83).

Following the reasoning of Calder, Insko, and Yandell (1974), Lutz and Swasy (1977) also suggested the expectancy-value approach. "...Subjects' free responses would be recast into belief statements of the form used by Fishbein and his associates...By allowing the subject to express his/her degree of agreement with the various thoughts that become salient, a more precise estimate of belief strength (B_i) can be obtained. In a similar vein, subjects would be asked to provide a_i (goodness-badness) ratings for the beliefs that they first mention spontaneously" (p. 367).

It would seem that a useful extension of this idea would be to use an approach that considers protocols as an indication of linkages that have been made between the message and elements of the cognitive structure (e.g., episodes, concepts, etc.) of the recipient. These would be more general than beliefs linking an object to its attributes (as in the Fishbein A_{object} model) or an act to its consequences (as in the Fishbein A_{act} term of his extended model). Then likelihood of the linkage's existence and the evaluation of the node to which message content has been linked could be assessed. This would allow a truly expectancy (likelihood)-value (evaluation) representation of the resultant, underlying, propositional structure that is probably indicated in the response protocols.

Construct validity of response protocols:

Regardless of the categories to which protocol responses are assigned, a crucial question in conceptualizing the positioning of these

measures in the elaborative model of persuasion is: do protocol responses represent thoughts that occurred as the message recipient was processing the message? If they do, then they can be considered to be measures of elaboration (a processing measure). If they don't, then they must be considered to be derived from the cognitive structure extant after message presentation (a structural measure).

Reactivity of the listing task. One of the most serious problems of asking Ss to list or recite their thoughts is the reactive nature of the task. Both the number and the nature of the thoughts elicited are probably affected by the following factors surrounding the task: setting, instructions, listing time available, and the timing of elicitation.

A. Reactive effects of the setting. The presence of other Ss may influence the number and nature of listed responses. "It would not be unusual for some subjects to give nonverbal reactions to a counter-attitudinal communication. A subject might shake his head, purse his lips, or give other visual signs indicating his disagreement with the message" (Osterhouse and Brock 1970, p. 356).

Ss may also be influenced by the experimenter. The knowledge that the experimenter will see or hear the responses may influence the number and content of thoughts listed ("...e.g., it's time to play the devil's advocate; I must be a critical, thoughtful person; etc..." (Roberts and Maccoby 1973, p. 301). "...Subjects may interpret the experimenter's request to produce counterarguments as a demand to justify retaining their original position" (Miller and Baron 1973, p. 108; see also Wright 1974a, p. 229).

B. Reactive effects of instructions. The instructions given to the Ss any time prior to listing should be structured to avoid influencing

the number or type of responses listed. There should also be no instructions prior to the message indicating that the study is examining thoughts that the S may have while listening to (reading, viewing) the communication. Such instructions could focus unnatural attention on their responses and change the number and nature of the thoughts listed.

Instructions given immediately prior to the listing task should contain no hint that the directionality or other aspect of the responses is of primary interest to the researcher. Unfortunately, Cullen (reported in Greenwald 1968) has had a lasting, negative effect on thought listing research in this regard. His instructions had numerous references to thoughts that may be "...favorable to one or the other viewpoint" (Greenwald 1968, p. 157). Instructions used by Calder, Insko, and Yandell (1974) and by Insko, Turnbull, and Yandell (1974) contained similar prompting. Petty, Wells, and Brock (1976) and Cialdini et al. (1976) both used instructions which included the following: "You might have ideas all in favor of the message you heard, all opposed, or a mixture of the two. Any of these is fine" (see also Petty and Cacioppo 1977, p. 651). While these instructions may generate thought listings that are easily coded into favorable and opposing categories, they certainly threaten the construct validity of the measures since they may engender evaluative responses where otherwise none may have occurred.

Another impact of the instructions on the validity of the protocol is whether the S is asked to refer to the communicative episode when performing the listing task. For example, Wright (1973a) asked Ss to list thoughts that "...had occurred to them during exposure" (p. 56), whereas Osterhouse and Brock (1970) simply asked for "thoughts and

ideas" about the topic (p. 346); and Petty, Wells, and Brock (1976) and Cialdini et al. (1976) asked Ss to "...write down the first idea that occurs to you on the message you heard..." (Petty 1977, emphasis added). Lack of explicit statements in the instructions to refer to the communicative episode raises questions about the researchers committment to the processing, rather than structural role of the statements produced. Yet, Osterhouse and Brock profess: "A crucial assumption of these two studies has been that the postcommunication assessment of counterarguments reflects prior counterarguing behavior during communication reception" (1970, p. 353).

C. Reactive effects of listing time. When Ss are given an indefinite period of time in which to list their thoughts, the chances of their writing thoughts that occurred to them during the listing task increases. Miller and Baron (1973) and Wright (1973a) suggest limiting response time as a means of reducing these spurious thoughts. "The essential idea is to allow sufficient time for complete recording of the honestly spontaneous thoughts, while reducing the probability that purely reactive thoughts are listed" (Wright 1973a, p. 56).

D. Reactive effects of timing. As discussed earlier, some studies elicited thoughts after other measures. When this occurs, Ss may try to maintain consistency between the earlier measures and the thoughts that they list. "In other words, scrutinizing a person's attitude may particularly predispose him to produce arguments that support his own position" (Miller and Baron 1973, p. 108).

In some cases, thoughts may be produced as rationalizations for the position that the person holds. Persuasion may lead to thought production, rather than message-relevant thoughts (which may be reported in

thought-listings) resulting in persuasion. (See Roberts and Maccoby 1973, p. 279.)

There may be consistency effects in reporting thoughts between types of responses. "...Expressions of source derogation may be the result of successful content-oriented counterargument..." (Miller and Baron 1973, p. 104).

Reliability and construct validity. Ironically, a study which examined the reliability of the thought-listing technique provides evidence against the validity of the procedure for recording thoughts that occurred during the communicative episode. Wright (see 1974a, p. 244 and 1975, p. 8) compared thoughts that were listed immediately after the communication with those collected from the same Ss two days later. "...71% of the total immediate thoughts reappeared identically in the delayed protocol..." (1974a, pp. 244-245). As Wright mentions, the act of recording may have contributed to the retention (1974a, p. 245). This, of course, also implies that the consistency across elicitation occasions may not be evidence of the reliability of the measure. However, it is also possible that both recordings simply report the result of the communicative episode, rather than the process of the episode itself.

Implications from cognitive psychological findings to the validity of protocol measures:

Thought listing is designed to be essentially a recall task. Ss are asked to recall thoughts that occurred to them during the communicative episode (depending on the instructions). The concepts of constructive and reconstructive memory processes discussed in Chapter 2 are applicable here. "Thoughts" may be easily generated inferentially from

the Ss' post-communication cognitive structures. They may report thoughts that they think are <u>likely to have occurred</u>, given their current position and their beliefs about the source. This effect would seem more likely if the S is somewhat familiar with the format and the content of the message (a likely situation with commercial messages).

The greater the delay between the communication and the listing task, the greater the likelihood that the listing will be derived from the underlying propositional structure that resulted from the interaction. "With time, perhaps, some of the experience may dissipate, thus making it difficult for the subject to discriminate what occurred in the episode from what the knowledge substructure activated can generate" (Brockway, Chmielewski, and Cofer 1974, p. 207).

Structural Criterion Variables:

If persuasion is viewed as shifting the cognitive structure of a message recipient to a position closer to that intended by the communicator, there are many dimensions along which such shifts may be measured. Within the perspective developed in Chapter 2, structural changes may involve establishing new linkages between concepts or events, changing the likelihood of linkages, changing the degree or amount of a characteristic that is linked to the topic, or changing the goodness or badness of concepts that have been associated with the topic or of the topic itself. (See Boyd, Ray, and Strong 1972; Wright 1973b; and Cohen 1974, for related discussions of advertising strategies.)

Which dimension is an appropriate measure of persuasive effectiveness depends on which change was intended by the communicator. <u>Attitude</u>, <u>desirability</u>, <u>liking</u>, <u>agreement</u>, <u>appropriate tuition level</u> are examples of structural measures that have been used. In some cases, the

criterion variable has been a single item and in other cases, it has been a composite of several items.

One of the most important issues regarding the use of structural measures in studies that assume an intervening role of elaboration is the problem of measuring both elaboration and structure in the same study. When both are measured, regardless of which comes first, there is always the possibility of convergence of the measures simply as a consistency artifact. This possible convergence is an important consideration whether a elaboration → post-cognitive-structure relationship is being examined or if a message → elaboration → post-cognitive-structure, a context → elaboration → post-cognitive-structure, or a prior cognitive structure → elaboration → post-cognitive-structure chain is being explored (see Figure 5). For studies examining the intervening role of elaboration in, say, a context → elaboration → post-cognitive-structure chain, it would probably be best to measure elaboration variables on one group of subjects in a context → elaboration study to establish that certain levels of elaboration accompany the levels of the contextual (situational) variable. Then, a context → post-cognitive-structure study could be performed on another group.

Another structural measure is recall. Recall merely assesses whether the individual remembers the relationships that were posited in a message, not whether the person believes them to be true. "The individual could memorize the content of the conclusion while his opinion remained unchanged" (Hovland, Janis, and Kelley 1953, p. 11). Recall has been measured using free recall, multiple choice and direct questions. (See Table 3, pp. 108-112).

Verbatim Recall Used as a Measure of a Communication's Effectiveness:

If the purpose of a communication is viewed as the transmitting of facts, verbatim recall could be a reasonable measure of success. However, message processing is not one of funneling information into storage in an unaltered form. It is an active process that results in a unique representation of the content that is compatible with the recipients' prior experiences and knowledge. If researchers insist on the reappearance of intact ideas or statements from the message in recall protocols, the scores developed in their coding schemes may misrepresent the retention of the gist of the message by the Ss. In fact, some statements that appear to be verbatim recall may be reconstructions (Cofer, Chmielewski, and Brockway 1976, p. 197).

Thought listing vs recall as a measure of cognitive integration:

Thought listing may be a better measure of message impact on the cognitive structure than requests to reproduce the text. Thought listing may include many of the associations that have been incorporated into the underlying propositional structure that is the S's cognitive representation of the gist of the message. Using thought listing, rather than recall, may provide the researcher with the richness of a view of the linkages resulting from message processing, rather than of an intact verbatim representation of the message which may remain isolated within the cognitive structure. The rich, elaborated representation provides the cognitive milieu to which the topic will be referred by the message recipient in future communicative and behavioral episodes.

Recall as a criterion measure in distraction studies:

When recall is used as a criterion variable in distraction studies, implications from cognitive psychology should be considered. First,

recall should decline when elaboration is restricted, as recall is related to the degree of elaboration. Second, the effects of distraction on recall should be more exaggerated when recall measures are delayed. "It follows from this interpretation that diverting processing capacity... should result in greater decrement in long-term that in short-term retention and, indeed there is good evidence that this is the case..." (Craik and Lockhart 1972, p. 670). None of the distractions studies reviewed here used delayed (other session) measures of recall!

Reported Results of Studies Examining the Relationships Portrayed in the Elaboration Model

Effects of Distraction on Post Cognitive Structure:

Table 5 presents the findings of 27 studies which examined the effect of distraction on structural variables. Twenty-six of these studies reported the relationship between distraction and some criterion that would reflect the position held by the message recipient (attitude, persuasibility, agreement, etc.). Nine of these indicated a significant main effect of distraction on the position-relevant measure. Of these 9, 4 reported a direct relationship, 2 an inverse relationship, and 2 suggested possible curvilinear relationships between distraction and the structural variable. With which explanation of distraction effects are these findings consistent? The thought-disruption hypothesis (which is essentially the same as that generated by the elaboration model) would predict greater attitude-change (acceptance) with increased distraction, as would the effort hypothesis. The studies by Festinger and Maccoby (1964), Osterhouse and Brock (1970), Keating and Latane (1972), and Keating and Brock (1974) all support these hypotheses. In each of these studies where recall was measured, it was unaffected by distraction.

Table 5. Main Effects of Distraction on Structural Variables.

Authors and Year	Criterion	Significance	Direction of Relationship
Festinger and Maccoby 1964			
Experiment 1	attitude	n.s	
Experiment 2	attitude	$p < 0.01$	direct
Freedman and Sears 1965	attitude	n.s.	
Rosenblatt 1966	attitude	$p < 0.025$	curvilinear (mod. highest)
Gardner 1966	desirability attitude recall	n.s. n.s. not reported	
Vohs and Garrett 1968	attitude	$p < 0.005$	inverse
Kiesler and Mathog 1968	communication effectiveness	*	
Haaland and Venkatesan 1968	attitude recall	$p < 0.05$ $p < 0.01$	inverse inverse
Silverman and Regula 1968	persuasibility	*	
Baron and Miller 1968	attitude	n.s.	
Osterhouse and Brock 1970			
Experiment 1	acceptance recall	n.s. n.s.	
Experiment 2	acceptance recall	$p < 0.05$ n.s.	direct

continued

Table 5 continued.

Authors and Year	Criterion	Significance	Direction of Relationship
Rule and Rehill 1970	agreement	*	
	retention	n.s.	
Gardner 1970	desirability	n.s.	
	attitude	n.s.	
	recall	$p < 0.001$	
Zimbardo et al. 1970			
Experiment 1	recall	$p < 0.002$	inverse
	attitude	n.s.	
Experiment 2	recall	$p < 0.01$	inverse**
	agreement	n.s.	
	evaluation of benefits	n.s.	
Experiment 3	recall	$p < 0.05$	inverse
	agreement	$p < 0.05$	curvilinear (message-set>)
	% pro speaker	$p < 0.005$	curvilinear (message-set>)
Keating and Latane 1972	attitude	$p < 0.01$	intermitent> continuous
Bither 1972	attitude	n.s.	
	recall	$p < 0.001$	inverse
Bither and Wright 1973	agreement	n.s.	
	recall	$p < 0.05$	inverse
Keating and Brock 1974	agreement	$p < 0.001$	direct
	appropriate tuition	$p < 0.001$	curvilinear?
	recall	n.s.	
Insko, Turnbull and Yandell 1974	agreement	$p < 0.001$	curvilinear?
	appropriate tuition	$p < 0.001$	curvilinear?
	recall	$p < 0.001$	inverse

continued

Table 5 continued.

Authors and Year	Criterion	Significance	Direction of Relationship
Petty, Wells and Brock 1976			
Experiment 1	agreement	*	
	appropriate	not reported	
	tuition	$p < 0.001$	inverse
	recall		
Experiment 2	agreement	*	
	appropriate	not reported	
	tuition	n.s.	
	recall		
Maccann Haslett 1976	agreement	n.s.	
	comprehension	n.s.	
Petty 1977			
Experiment 2	attitude	*	
	appropriate	not reported	
	tuition		

* involved in an interaction

** assuming that message-only is the lowest distraction condition

For the Festinger and Maccoby (1964) findings, the affect hypothesis is equally as good as the disruption hypothesis, since a humorous film was used as the distractor and no manipulation check was performed. The Keating and Latane (1972) study varies signal quality which may not induce distraction (no manipulation check was performed). The Osterhouse and Brock (1970) and Keating and Brock (1974) studies both involved a flashing-light-monitoring task. This task is quantitative, externally paced, requires a response, and has varying intervals between flashes (see Table 1, pp. 79-82). As such, it has face validity as a task which would require allocation of the limited processing capacity. Manipulation checks for both studies also indicated successful manipulation of distraction (see Table 2, pp. 89-92). These two studies support the thought-disruption hypothesis.

The studies by Vohs and Garret (1968) and by Haaland and Venkateson (1968) both indicated an inverse relationship between distraction and persuasion, which is compatable with the learning hypothesis. Although no manipulation check was performed, the fact that an overt response was required and the tasks were resource limited lends face validity to the manipulations. Although recall was not measured in the Vohs and Garret studies (1968), an earlier study by Vohs (1964) indicated that the tasks significantly reduced retention. Recall was also significantly reduced by distraction in the Haaland and Venkatesan (1968) study. It is interesting that the visual distraction in the Haaland and Venkatesan (1968) study (a humorous film) had the opposite effect to that found in the Festinger and Maccoby (1964) study.

In the studies where <u>recall was unaffected</u> and there was a direct relationship between distraction and attitude, the distraction task was

data limited (that is, a simple low-level response was required). In studies where recall was reduced by distraction and attitude was inversely related to distraction, the distraction task was resource limited (more extensive processing was required to respond to the distractor). These differential results suggest a possible curvilinear effect, with "mild" distraction reducing opposition to the message and higher levels reducing the transmission of the underlying propositions of the message, as well as any spontaneous, elaborative responses. Both sets of findings are in accord with the recall condition (specified in Osterhouse and Brock (1970) and clearly stated in Keating and Brock (1974, p. 302)) for which distraction will produce yielding to the message: "...the distracting task must interfere with counterargument without at the same time, affecting the reproducibility of the message".

Curvilinear effects were found in three studies. Rosenblatt (1966) found that the "moderate" condition was significantly more effective in inducing yielding. However, it is not clear what was manipulated in this experiment (there was not manipulation check) and the "moderate" label may be purely arbitrary.

The other studies which indicate a curvilinear effect are those by Zimbardo et al. (1970) and by Insko, Turnball and Yandell (1974). Concluding that there was a curvilinear effect in these studies is dependent upon viewing the message-set condition (which involved a concurrent task) as being a higher distraction condition than the message-only condition (which involved no other task). "The message-set group agreed significantly more with the speaker than either the message-only or number-set groups..." (Zimbardo et al. 1970, p. 678). However, a review of the manipulation check performed in the Zimbardo et al. (1970) study

reveals that, although Ss in the message-only condition reported greater effort and greater allocated attention than the message-set groups, the differences were not significant for these two groups on either measure (see Table 2, pp. 89-92).

In the Insko, Turnbull, and Yandell (1974) study there was again no significant difference between the message-only and message-set groups on the effort manipulation check measure. Since their findings confirm those of Zimbardo et al. (1970) they suggest "...that a dissonance or attribution model cannot account for the facilitative effect of message-set upon attitude change. Message-set increased attitude change [over message-only], but did not alter reported effort" (p. 518). (The dissonance/attribution model would predict consistency between the level of agreement with the message and the effort expended to process it. It would be inconsistent for a person to have expended effort to process a discrepant message.) Recall was significantly reduced in the message-set condition as compared to the message-only condition. "It should be noted that these results do not agree with Osterhouse and Brock's generalization that distraction only facilitates persuasion where there is no interference with recall" (Insko, Turnbull, and Yandell 1974, p. 518).

The failure to achieve significant differences in attention or effort between the message-only and message-set conditions raises questions about the role of attention allocation in the persuasion process, since attitude was significantly greater for the message-set condition in both studies. It may be that self-report of attention and effort is not as sensitive as measures of attitude. These self-report measures may also lack construct validity as measures of actual, rather

that perceived, attention allocation. If they are valid measures, the elaboration model may need revision. Perhaps a future study utilizing quantitative, objective measures of distraction will resolve this issue.

Six studies report interactive effects on attitude measures. Kiesler and Mathog (1968) report an interaction between credibility and distraction. "High Interference raised communication effectiveness under High Credibility but not under Low Credibility... High Interference also lowered communication effectiveness under Low Credibility" (Kiesler and Mathog 1968, p. 1130). They conclude that their findings favor the disruption, rather than the effort hypothesis. "Our case, however, rests upon an assumption not directly tested: that counter-arguing is greater under high credibility than low credibility" (Kiesler and Mathog 1968, p. 1131). Later studies found an inverse relationship between counterarguing and credibility (see Table 7, pp. 153-159).

Silverman and Regula (1968) found an interaction between intentionality and distraction. The effect of distraction was greater when Ss were told that static on the tape was a part of the experiment. However, this study may not show that distraction effects only occur in reactive settings because they may not have manipulated distraction, but rather simply varied signal quality. In fact, Ss may have allocated <u>more</u> attention to the message when static was present in an effort to understand the message (especially if they were told that it was part of the experiment).

Rule and Rehill (1970) found that distraction only affected agreement for Ss in the success condition (Ss were told that they had successfully completed a difficult maze task before other Ss). However, it was not established that distraction had been manipulated. It seems

that the effects could as easily be attributed to the reactive nature of the experiment. Ss in the success condition who hear an unexpected series of 22 different sound effects could view it as a test of their abilities and indicate agreement to show that they had successfully understood the message. Also, Ss may have concentrated harder on their reading to shut out the sounds to show the experimenter that they could also be successful at this task, so that the findings may reflect a learning effect.

A message-distraction interaction was found in Petty, Wells, and Brock (1976 Experiments 1 and 2; experiment 2 is also described in Petty 1977)). For both experiments, when the message utilized ineffectual arguments, increasing acceptance accompanied increasing distraction. The opposite effect was found when the message included difficult-to-counterargue arguments. In those conditions, acceptance decreased with increasing distraction (however, this effect was non-significant in Experiment 2). Petty, Wells, and Brock cite their findings as supporting the distraction hypothesis, rather than the effort hypothesis: "The thought disruption hypothesis allows the specific prediction that distraction is most likely to lead to enhanced persuasion when a message presents poor refutation and counterargumentation) and to reduced persuasion when a message presents very good arguments (i.e., arguments that likely to elicit favorable thoughts). The effort justification hypothesis fails because it is unable to make such differential predictions. (1976, p. 883). Unfortunately, the question that they asked to measure effort ("To what extent did you find the task effortful?" (p. 878)) was ambiguous as to what "task" it referred to. Petty, Wells and Brock (1976) also report a result

that is contrary to the recall condition stated by Osterhouse and Brock (1970) (as did Insko, Turnbull, and Yandell (1974)). The highest level of distraction significantly reduced recall, yet had the highest acceptance level.

All of the studies that examined the effect of distraction on recall and reported findings (10 of 17 studies) found an inverse relationship. This is consistent with the elaboration model. The delayed effect should be greater than the short-term effect. Unfortunately, delayed recall measures were not taken.

Effects of Distraction on Measures of Elaboration:

Table 6 presents the findings of studies which examined the effect of distraction on elaborative responses. All of the significant findings indicate an inverse relationship between distraction and elaboration. However, only indices involving counterarguments were significantly related to distractions. When positive responses or the number of words or thoughts were the criterion variable, the findings were non-significant. However, there were no significant main effects for any predictor variables in those studies.

Petty, Wells, and Brock (1976, see also Petty 1977) found an interactive effect of distraction and message counterarguability on favorable thoughts (proarguments). They found that distraction had no effect on proargument production for the easy-to-counterargue message. However, increasing distraction significantly reduced the production of proarguments for the difficult-to-counterargue message. This finding is compatable with the elaboration model and the thought disruption hypothesis, as distraction had the greatest negative effect in conditions where proarguments were most likely.

Table 6. Main Effects of Distraction on Measures of Elaboration (B-D studies) (CA = counterargument).

Authors and Year	Criterion	Significance	Direction of Relationship
Baron and Miller 1969	number of words number of thoughts	n.s. n.s.	
Osterhouse and Brock 1970 Experiment 1	number of CA	$p < 0.05$	inverse
Experiment 2	number of CA	$p < 0.03$	inverse
Zimbardo et al. 1970 Experiment 2	percent of time spent CA	$p < 0.001$	inverse
Experiment 3	percent of time spent CA	$p < 0.005$	inverse
Keating and Brock 1974	number of CA	$p < 0.001$	inverse
Insko, Turnbull and Yandell 1974	number of CA percent of time spent CA	$p < 0.05$ n.s.	inverse
Petty, Wells and Brock 1976 Experiment 1	number of CA favorable thoughts	$p < 0.05$ n.s.	inverse
Experiment 2	number of CA favorable thoughts	n.s. *	
Maccann Haslett 1976	number of CA	$p < 0.001$	inverse
Petty 1977 Experiment 2	number of CA number of pro-arguments	n.s. *	

* involved in an interaction

The study by Roberts and Maccoby (1973) reported that Ss who were to write their responses, rather than just think during message presentation (and/or during pauses) produced significantly fewer counterarguments in post-message protocols. Even thought the authors did not position this variable as a distractor, it seems to meet the requirements.

The Effect of Predictor Variables (Other than Distraction) on Measures of Elaboration:

The effects of non-distraction variables on elaboration are presented in Table 7. Anticipatory (prior to message) counterarguments increased with: (1) increasing discrepancy (extremity) of the message; (2) presentation of an example counterargument; (3) forewarning; (4) increased time since forewarning; and (5) when subjects were asked to list topic related rather than actual thoughts. Anticipatory thoughts (and words) decreased with increasing credibility of the spokesperson (Baron and Miller 1969). Anticipatory favorable thoughts increased with: (1) forewarning; (2) time since forewarning; and (3) when subjects were asked to list topic related rather than actual thoughts.

Ss were significantly more likely to counterargue following message presentation when: (1) there were more arguments in the message, (2) they had more opportunity (time), (3) they could just think, rather than write during message presentation, (4) they were more confident (both socially and in their ability to process information), (5) they were designated as opinion leaders, (6) the message used weak arguments, (7) they were forewarned, (8) they were told to write statements related to the topic in their protocols, (9) the message source lacked credibility and (10) the message was counterattitudinal.

Table 7. Effects (other than distraction) on Measures of Elaboration (N = number, CA = counter-arguments, SD = source derogation, FT = favorable thoughts, NT = neutral thoughts and SA = support argument).

Author and Year	Predictor	Criterion	Significance	Direction of Relationship
Krugman 1966	medium	N of connections	not reported	
	product	N of connections	not reported	
	editorial environment	N of connections	not reported	
	instructional set	N of connections	not reported	
Brock 1967	discrepancy	N of CA (anticipatory)	$p < 0.001$	direct
	persuasive intent	N of CA	n.s.	
	priming	N of CA	$p < 0.001$	direct
Baron and Miller 1969	credibility	N of thoughts (anticipatory)	$p < 0.02$	inverse
		N of words	$p < 0.0025$	inverse

continued

Table 7 continued.

Author and Year	Predictor	Criterion	Significance	Direction of Relationship
Cook 1969 Experiment 1	competence	N of CA	$p < 0.05$	inverse
	reception	N of CA	$p < 0.001$	direct
Experiment 2	competence	N of CA	*	
	reception	N of CA	*	
Osterhouse and Brock 1970 Experiment 2	threat	N of CA	*	
	degree of influence	N of CA	*	
Leavitt, Wadell and Wells 1970	time since eating	Personal Product Response	n.s. (commercial A) $p < 0.05$ (commercial B)	direct

continued

Table 7 continued.

Author and Year	Predictor	Criterion	Significance	Direction of Relationship
Maccoby and Roberts 1972		A. Measures taken after message presentation:		
	opportunity	N comments	n.s.	
		N supportive comments	n.s.	
		N counter comments	$p < 0.05$	"stops">continuous
		intensity (algebraic sum)	$p < 0.01$	"stops"<continuous
	think vs write during presentation	N comments	n.s.	
		N supportive comments	n.s.	
		N counter comments	$p < 0.05$	think>write
		intensity (algebraic sum)	$p < 0.05$	think<write
		B. Measures taken during message presentation:		
	opportunity	N comments	$p < 0.05$	"stop">continuous
		N supportive comments	n.s.	
		N counter comments	$p < 0.001$	"stop">continuous
		intensity (algebraic sum)	n.s.	
Roberts and Maccoby 1972	essentially a reworking of the experimental data presented in Maccoby and Roberts 1972			

continued

Table 7 continued.

Author and Year	Predictor	Criterion	Significance	Direction of Relationship
Wright 1974b	media	total cognitive activity	$p < 0.01$	print>broadcast
		N of CA	* †	
		N of SD	$p < 0.05$	print>broadcast
		N of SA	$p < 0.001$	print>broadcast
	content involvement	total cognitive activity	n.s.	
		N of CA	* †	
		N of SD	$P < 0.05$	low involvement> high involvement
		N of SA	n.s.	
	†media x involvement	N of CA	$p < 0.05$	print: high involvement>low involvement
Insko, Turnbull and Yandell 1974	internal vs external Ss	no significant effects		
	order of attitude-thought listing	no significant effects		
	thin-normal-fat Ss	no significant effects		

continued

Table 7 continued.

Author and Year	Predictor	Criterion	Significance	Direction of Relationship
Calder, Insko, and Yandell 1974	N of prosecution arguments	N of thoughts	n.s.	
		N of defense thoughts	$p < 0.01$	inverse
		N of prosecution thoughts	$p < 0.01$	direct
		prosecution minus defense	n.s.	
	N of defense arguments	N of thoughts	n.s.	
		N of defense thoughts	*	
		N of prosecution thoughts	*	
		prosecution minus defense	$p < 0.01$	inverse
	time of assessment	N of thoughts	*	
		N of defense thoughts	*	
		N of prosecution thoughts	*	
		prosecution minus defense	n.s.	
Wright 1975	Generalized Social Confidence (audio/ high involvement)	N of CA	$p < 0.01$	direct
		N of SA	n.s.	
		N of SD	n.s.	
	Generalized Social Confidence (audio/ low involvement)	N of CA	n.s.	
		N of SA	n.s.	
		N of SD	n.s.	

continued

Table 7 continued.

Author and Year	Predictor	Criterion	Significance	Direction of Relationship
Wright 1975 continued	Information Processing Confidence (audio/high involvement)	N of CA	n.s.	
		N of SA	n.s.	
		N of SD	n.s.	
	Information Processing Confidence (audio/low involvement)	N of CA	$p < 0.01$	direct
		N of SA	n.s.	
		N of SD	n.s.	
	Peer-designated opinion leader (audio/high involvement)	N of CA	$p < 0.01$	direct
		N of SA	n.s.	
		N of SD	n.s.	
	Peer-designated opinion leader (audio/low involvement)	N of CA	n.s.	
		N of SA	n.s.	
		N of SD	$p < 0.05$	direct
Petty, Wells and Brock 1976 Experiment 1	message counter-arguability	N of CA (manipulation check)	$p < 0.02$	direct
		N favorable thoughts	n.s.	

continued

Table 7 continued.

Author and Year	Predictor	Criterion	Significance	Direction of Relationship
Petty, Wells and Brock 1976 Experiment 2	message counter-arguability	N of CA (manipulation check)	$p < 0.0001$	direct
		N favorable thoughts	$p < 0.001$	inverse
Cialdini et al. 1976	high-relevance/immediate discussion versus other 3 conditions	N supportive thougths	$p < 0.011$	high relevance/immediate discussion < other conditions
		N other thoughts	n.s.	
Petty 1977 Experiment 1	forewarning	N of CA	$p < 0.009$	warned > unwarned
		N of pro-arguments	n.s.	
	instructions (topic versus actual thoughts)	N of CA	$p < 0.001$	topic > actual
		N of pro-arguments	n.s.	
Experiment 2	message quality	N of CA	$p < 0.001$	inverse
		N of pro-arguments	$p < 0.001$	direct
Experiment 3	body posture	N of CA	n.s.	
		N of pro-arguments	n.s.	

continued

Table 7 continued.

Author and Year	Predictor	Criterion	Significance	Direction of Relationship
Petty and Cacioppo 1977				
Experiment 1	immediate vs no warning	N of CA	n.s.	
		N of FT	n.s.	
	5.25 minute warning vs no warning	N of CA	$p < 0.001$	warned>unwarned
		N of FT	$p < 0.01$	warned>unwarned
	5.25 minute warning vs immediate warning	N of CA	$p < 0.001$	5.25 min>immediate
		N of FT	$p < 0.01$	5.25 min>immediate
Experiment 2	warning vs no warning	N of CA	$p < 0.009$	warned>unwarned
		N of NT	$p < 0.02$	unwarned>warned
		N of FT	n.s.	
	instructions (topic thoughts vs actual thoughts)	N of CA	$p < 0.001$	topic>actual
		N of FT	$p < 0.001$	topic>actual
		N of NT	$p < 0.001$	actual>topic
Cacioppo, Sandman and Walker 1978	heartrate	N of CA	$p < 0.10$	accelerated>decelerated
		N of FT	n.s.	
		N of NT	n.s.	

continued

Table 7 continued.

Author and Year	Predictor	Criterion	Significance	Direction of Relationship
Cacioppo, Sandman and Walker 1978 continued	message persuasiveness	N of CA	$p < 0.05$	low-persuasive> high persuasiveness
Cacioppo and Petty 1979	counterattitudinal vs proattitudinal message	N of CA	$p < 0.001$	counterattitudinal> proattitudinal
		N of neutral/ irrelevant	$p < 0.05$	proattitudinal> counterattitudinal
	number of message repetitions	N of CA	$p < 0.001$	quadratic trend
		N of neutral/ ireelevant	$p < 0.02$	linear trend
Petty and Cacioppo 1979 Experiment 1	message orientation (counter attitudinal vs proattitudinal)	N of CA	* ††	
		N of FT	* ††	
	issue involvement	N of CA	* ††	
		N of FT	* ††	

continued

Table 7 continued.

Author and Year	Predictor	Criterion	Significance	Direction of Relationship
Petty and Cacioppo 1979 Experiment 1 continued	††orientation X issue involvement			
	1. high issue involvement:	N of CA	$p < 0.05$	counterattitudinal> proattitudinal
	counterattitudinal vs proattitudinal	N of FT	$p < 0.05$	proattitudinal> counterattitudinal
	2. low issue involvement: counterattitudinal vs proattitudinal	N of CA	n.s.	
		N of FT	n.s.	
Experiment 2	argument quality (weak vs strong arguments)	N of CA	* †††	
		N of FT	* †††	
	involvement	N of CA	* †††	
		N of FT	* †††	
	†††argument quality X involvement			
	1. high involvement: weak vs strong arguments	N of CA	$p < 0.05$	weak>strong
		N of FT	$p < 0.05$	strong>weak
	2. low involvement: weak vs strong arguments	N of CA	n.s.	
		N of FT	n.s.	

continued

Table 7 continued.

Author and Year	Predictor	Criterion	Significance	Direction of Relationship
Petty and Cacioppo 1979 Experiment 2 continued	3. weak arguments: high vs low involvement	N of CA	$p < 0.05$	high>low
		N of FT	n.s.	
	4. strong arguments: high vs low involvement	N of CA	n.s.	
		N of FT	$p < 0.05$	high>low

* involved in an interaction
** manipulation check indicated that this variable was not successfully manipulated
†,††,††† refer to interaction reported directly below

Supportive comments increased when (1) media was print, rather than broadcast and (2) message quality increased.

The frequency of occurrence of a Personal Product Response in Ss' protocols was significantly related to time-since-eating (direct relationship) for one commercial, but not for another. (The authors reported one, combined, significant Chi Square analysis. However, examination of their data reveals the differential results.)

Source derogation increased when the message was presented in print, rather than broadcast; when the Ss were in the low, rather than high involvement condition; and with increases in peer-designated topical opinion leadership (for audio/low involvement settings).

Protocol statements coded as "defense" thoughts decreased as the number of "prosecution" arguments presented increased. "Prosecution" thoughts increased with increasing "prosecution" arguments and decreasing "defense" arguments.

Several studies revealed interactive effects on measures of elaboration. In Experiment 2, Cook (1969) found that when there were few arguments ("low reception"--2 arguments), competence (credibility) did not affect reported counterarguments, but for "high reception" (10 arguments), the number of counterarguments was inversely related to competence. When Ss can only respond to two arguments there can be little variance in their responses across competence conditions. There were always more counterarguments observed in this experiment for the high reception conditions, regardless of the competence of the source.

Osterhouse and Brock (1970, Experiment 2) found an interaction between threat and perceived influence on counterargument production. "When exposed to a strongly threatening communication high-perceived-

influence subjects counterargued more than did low-perceived-influence subjects. No such tendency was found under the mild-threat condition" (p. 352). As noted in the discussion of predictor variables, their manipulation of threat may have also manipulated credibility. However, the findings here are contrary to other reported effects of credibility on counterarguing.

Wright (1974b) found that when the message was presented in print, rather than broadcast, high-content-involvement Ss counterargued more than low-content-involvement Ss. He argued that the effect occurs for print since this is self-paced, whereas broadcast messages do not provide enough time for content-based counterarguments.

Calder, Insko and Yandell (1974) reported a time-of-assessment by number-of-defense-arguments interaction on both the number of defense and prosecution thoughts. However, the number of defense arguments always had an inverse effect on prosecution thoughts and a direct effect on defense thoughts. (The reader is referred to Calder, Insko, and Yandell (1974, pp. 77-78) for a discussion of the interaction.)

In Experiment 1, Petty and Cacioppo (1979) found an interaction between message orientation and issue involvement on both the number of counterarguments and the number of favorable thoughts. With high issue involvement counterattitudinal messages produced more counterarguments and fewer favorable thoughts than did proattitudinal messages. In Experiment 2 they found an interaction between issue involvement and argument quality on both the number of counterarguments and the number of favorable thoughts. With high issue involvement strong arguments produced fewer counterarguments and more favorable thoughts than did weak arguments. High issue involvement produced more counterarguments

than did low issue involvement when weak arguments were used and more favorable thoughts when strong arguments were used.

Relationships Between Elaborative and Structural Responses:

Table 8 presents the reported relationships between structural and elaborative measures. Of the eleven studies that reported examination of these relationships, only one found no significant relationships. The results show that the two types of measures are compatable. As counterarguments increase, structural responses that indicate message acceptance (attitude, agreement, convincingness, buying intentions) decrease. As supportive elaborative responses increase, acceptance increases. As prosecution thoughts increase relative to defense thoughts, certainty of guilt increases.

These findings are important for two reasons. First, if elaborative measures are measures of elaborative activity, then the elaboration model would fail if these findings had not occurred. Second, even if elaborative measures lack construct validity and are instead reconstructions based on post-message cognitive structure, they may be surrogates for structural measures. When the instructions are neutral, elaborative measures have the advantage over structural measures in their lack of reactivity. "One of the beauties of the thought-monitoring approach is its indirect manner of extracting the response variables of interest. When the subject is asked to report all thoughts that occurred, he does so in ignorance of what type of response the researcher will be singling out for examination" (Wright 1974a, p 235).

The use of elaborative measures as surrogates for structural measures was suggested by Greenwald (1968) in his review of Cullen's findings: "The correlations with post-test opinion for the recipient-

Table 8. Elaboration-Structural Relationships

Author and Year	Elaboration Measure	Structural Measure	Significance	Direction of Relationship
Brock 1977	pre-message counterarguments	1. post-message agreement	$p < 0.01$	inverse
		2. convincingness	$p < 0.05$	inverse
Grenwald 1968	directional content index			
A. Cullen	1. externally originated	opinion (post-test)	n.s.	
	2. recipient-modified	opinion (post-test)	n.s.	
	3. recipient-generated	opinion (post-test)	$p < 0.01$	direct
B. Love	intermittent (during cognitive reactions)	recall	n.s.	
Osterhouse and Brock 1970 Experiment 1	counterarguments	acceptance	$p < 0.01$	inverse
Experiment 2	counterarguments	acceptance	$p < 0.01$	inverse
Wright 1973a	1. $\Sigma_i CA_i$	1. convincingness (A_1)	$p < 0.001$	inverse
		2. attitude (A_2)	$p < 0.001$	inverse
		3. buying intention (BI)	$p < 0.001$	inverse

continued

Table 8 continued.

Author and Year	Elaboration Measure	Structural Measure	Significance	Direction of Relationship
Calder, Insko, and Turnbull 1974	1. external (number prosecution-number defense)	belief (certainty of innocence-guilt; 1-9 scale; 9 = "certainty of guilt")	$p < 0.01$	direct
	2. recipient-modified (prosecution-defense)		$p < 0.01$	
	3. recipient-generated (prosecution-defense)		$p < 0.05$	direct
Keating and Brock 1974	number of counter-arguments*	agreement	$p < 0.01$	inverse
Petty 1977 Experiment 3	number of counter-arguments*	agreement	$p < 0.05$	inverse
	number of pro-arguments*	agreement	$p < 0.05$	direct
Cacioppo, Sandman and Walker 1978	number of counter-arguments*	agreement	$p < 0.10$	inverse
Cacioppo and Petty 1979	number of counter-arguments*	agreement	$p < 0.05$	inverse
		recall	$p < 0.05$	direct
	number of favorable thoughts*	agreement	$p < 0.05$	direct
		recall	n.s.	

continued

Table 8 continued.

Author and Year	Elaboration Measure	Structural Measure	Significance	Direction of Relationship
Wright 1973a continued	2. $\Sigma_i SA_i - \Sigma_i CA_i$	1. A_1	$p < 0.001$	direct
		2. A_2	$p < 0.001$	direct
		3. BI	$p < 0.001$	direct
	3. $\Sigma_i SA_i - \Sigma_i CA_i - \Sigma_i SD_i$	1. A_1	$p < 0.001$	direct
		2. A_2	$p < 0.001$	direct
		3. BI	$p < 0.001$	direct
	4. $\Sigma_i W_i SA_i - \Sigma_i W_i CA_i$	1. A_1	$p < 0.001$	direct
		2. A_2	$p < 0.001$	direct
		3. BI	$p < 0.001$	direct
	5. $\Sigma_i W_i SA_i - \Sigma_i W_i CA_i - \Sigma_i W_i SD_i$	1. A_1	$p < 0.001$	direct
		2. A_2	$p < 0.001$	direct
		3. BI	$p < 0.001$	direct
	6. Advertising-originated	1. A_1	$p < 0.05$	direct
		2. A_2	n.s.	
		3. BI	n.s.	
	7. recipient-modified	1. A_1	$p < 0.01$	direct
		2. A_2	$p < 0.01$	direct
		3. BI	$p < 0.01$	direct
	8. recipient-generated	1. A_1	$p < 0.01$	direct
		2. A_2	$p < 0.01$	direct
		3. BI	$p < 0.01$	direct

continued

Table 8 continued.

Author and Year	Elaboration Measure	Structural Measure	Significance	Direction of Relationship
Petty and Cacioppo 1979 Experiment 2	number of counter-arguments*:			
	(1) low involvement	attitude index	n.s.	
	(2) high involvement	recall	n.s.	
	number of favorable thoughts*:			
	(1) low involvement	attitude index	$p < 0.05$	direct
		recall	n.s.	
	(2) high involvement	attitude index	$p < 0.001$	direct
		recall	n.s.	

* followed structural measure

generated category index based on all thoughts combined were sufficiently high to suggest that the thought-listing procedure used in the present experiment might be useful as a measure of opinion" (p. 161). The cummulative evidence supports his position.

One of the studies, Cacioppo and Petty (1979) reported a significant, direct relationship between the number of counterarguments and recall. This is important as the elaboration model suggests that recall should increase with increasing elaboration.

Attempts to Establish the Intervening Nature of Elaborative Responses:

Analysis of covariance has been used in six studies in an attempt to establish whether reported elaborative responses represent an intervening process in persuasion (Cook 1969; Osterhouse and Brock 1970; Roberts and Maccoby 1973; Insko, Turnbull, and Yandell 1974; Petty and Cacioppo 1977; and Cacioppo and Petty 1979).

Cook (1969, Experiment 1) reported that using counterarguments as a covariate with attitude as the criterion reduced F values more than the use of Ss' reported trust in or powerfulness of the arguments in the message. He concluded that: "Within high reception conditions, then, counterarguing may be the strongest cause of the difference attributable to competence" (p. 351). He did not try using counterarguments as the criterion and attitude as the covariate. If both attitude and counterarguing reflect the Ss' cognitive structures, using counterarguments as a covariate would reduce the F values whether or not it reflects an intervening process. Cook also reported: "Unfortunately, Blalock's (1961) partial correlation techniques, using rated expertise, counterarguing, and final attitude, failed to discriminate whether counterarguing caused final attitude or whether final attitude caused counterarguing" (p. 355).

In their study of the effects of distraction, Osterhouse and Brock (1970) used counterarguments as the covariate and acceptance as the criterion. The F value was reduced from 2.38 to 0.82. But, the <u>initial</u> F value (2.38) was non-significant. However, the previously significant linear trend ($F = 3.41$, $p < 0.05$) was reduced to non-significance ($F = 0.85$).

In Experiment 2, Osterhouse and Brock reported that again, with acceptance as the criterion and counterarguments as the covariate, the F value for the distraction effect was reduced (in this case from $F = 3.0$, $p < 0.05$ to a non-significant $F = 1.02$). However, just as with Cook (1969), acceptance was not used as the covariate. (See Lutz and Swasey 1977, p. 367).

Roberts and Maccoby (1973) reported changes in F values for the effect of opportunity on opinion with each of their many categories of elaborative response as covariates. When responses from protocols generated after reception were used as covariates, none reduced the F value to non-significance. When responses from protocols generated during reception were used as covariates, three of five "subject generated" categories reduced the F values to non-significant levels. However, the significance level associated with these reduced F values was $p < 0.10$.

Based on the reduced F values and on their timing of the protocol elicitation (during "broadcast" message presentation), Roberts and Maccoby (1973) argue that the reported elaborative responses must reflect an intervening (mediating) role in the persuasion process. Cook's (1969) Ss also had the opportunity to write their protocols while reading the message. Osterhouse and Brock (1970) elicited protocols after intervening structural measures.

Roberts and Maccoby (1973) also failed to use opinion as a covariate with protocol responses as the criterion variable. So, again, the reported elaborations and the structural variables may simply be alternative ways for Ss to reveal their cognitive structures.

Insko, Turnbull, and Yandell (1974) also used covariance analysis, albeit, in only one direction. When they used counterarguments as a covariate, the F value for the effect of perceived task on a combined structural variable was reduced from $F = 3.55$ ($p < 0.04$) to $F = 0.37$. Using communication-favorable thoughts as a covariate reduced the F value to 0.96.

Insko, Turnbull, and Yandell (1974) proposed that communication-favorable thoughts temporally precede counterarguments. When they used communication-favorable thoughts as a covariate, the F value for the effect of perceived task on counterargument production dropped from $F = 9.02$ ($p < 0.01$) to $F = 2.66$ ($p < 0.11$). Again, only one direction was tested.

Petty and Cacioppo (1977) used covariance analysis both with protocol responses and with one of their structural measures (attitude) as the covariate. When they examined the effect of forewarning on attitude with the number of counterarguments as the covariate, the F value was reduced from $F = 5.30$ ($p < 0.02$) to a non-significant $F = 0.26$. When the number of favorable thoughts was used as the covariate the F value <u>increased</u> to 12.60.

When attitude was used as the covariate and the number of counterarguments as the criterion, the F value for the effect of forewarning was reduced from $F = 19.03$ to $F = 12.66$, but both are significant. They do not report the results of a test using attitude as the criterion and the number of favorable thoughts as the covariate.

Although the findings of Petty and Cacioppo with regard to counter-arguments are consistent with a model which posits elaborative responses as a mediator of attitudes, their findings with regard to favorable thoughts are not. They also warn that "interpretation of these analyses should be cautious because it is unclear whether extreme departure from the ANCOVA assumptions can be tolerated" (p. 653).

Cacioppo and Petty (1979) also used analysis of covariance to examine the plausibility of reported elaborative responses as a mediator of structural changes. When agreement was used as the criterion, the use of both counterarguments and favorable thoughts reduced the F value for the effect of message repetition from $F = 3.04$ ($p < 0.05$) to $Fs < 1$. When agreement was used as the covariate with favorable thoughts as the criterion the F value for the effect of repetition was reduced from $F = 2.36$ ($p < 0.01$) to $F = 1.78$ ($p < 0.20$). The F value for the effect of repetition on counterarguing was reduced from $F = 1.57$ ($p < 0.20$) to $F = 0.68$. Since the F values were reduced both when reported elaborative responses were used as covariates and when agreement was used as a covariate, the results are ambiguous support for reported elaborative responses serving as a mediator of structural change.

In summary, all six studies showed that the elaborative responses are related to structural responses. However, the causal nature of the relationship remains unsettled.

Conclusion

What can be said, then, about the elaborative model of persuasion? When various variables are introduced which should increase or decrease elaborative responses, protocol reports reflect effects that would be

expected if elaboration were, in fact, being measured. Tendencies observed in protocol reports are also reflected in structural responses. In sum, the circumstantial evidence points to the existence and intervening role of elaborative responses. Yet, because of experimental limitations, true elaborative responses may never have been recorded.

What, then, of the use of elaborative protocols? They remain a useful tool as both a criterion and as a predictive measure. However, it is still unclear exactly what protocols represent--reconstructions, accurate reports of elaborative episodes, or simply attempts by Ss to persuade the experimenter. What is measured probably varies from study to study--with different instructions, timing, and settings. Unless the construct validity of the measure has been established for a study, researchers should exhibit restraint in presenting verbal report data as being derived from measures of elaboration.

CHAPTER FOUR

SOURCES OF VARIANCE IN THE CRITERION VARIABLE: IMPLICATIONS AND CONTROL

The elaboration model of message processing incorporates many variables which are not directly observable (credibility, prior cognitive structure, etc.). These variables can be especially troublesome in an experimental setting. The experimenter would like to be able to attribute variations in the criterion variable(s) to changes in the manipulated variables. However, since many of the predictor variables are unobservable, steps must be taken to establish that they have, in fact, been manipulated. In addition, extraneous variables may intrude in the experimental setting and confound the attribution of causality. This chapter will provide guidelines for dealing with these and other experimental problems.

The chapter begins with an examination of the requirements for making causal inference. Following this is an examination of sources of variation in the criterion variable and the implications of various sources of variance to the causal inference process. Methods of controlling different sources of variation are then discussed. The chapter ends with a discussion of persistent sources of variance and their control: the experimenter, subjects, instructions and questionnaires.

Requirements for Valid Causal Inference

An experiment attempts to establish the plausibility of the statement $\Delta \underline{X} \rightarrow \Delta \underline{Y}$ (a difference in $\underline{X}$ (the independent or predictor variable) leads to a difference in $\underline{Y}$ (the dependent or criterion

variable)). Several requirements must have been met before this plausibility may be claimed. It must have been established that: (1) that there was indeed a change in X; (2) that there was indeed a change in Y; (3) that ΔX led to ΔY and not the reverse; (4) that the particular exemplar, X´, was representative of the theoretical independent variable of interest (i.e., that it has construct validity); (5) that the particular exemplar Y´, is representative of the theoretical criterion variable of interest; (6) that there were no variables present in the experiment other than X whose change could have plausibly led to ΔY.

For the statement ΔX → ΔY to be useful in explaining a phenomenon, two additional conditions should be met: (7) the generality of the statement should be established (i.e., the contexts beyond the experiment proper for which the statement may be expected to hold should be determined (external and ecological validity)); and (8) the statement should not be reduced to a tautology by the use of an X´ and Y´ that are both exemplars of the same underlying theoretical variable.

Requirement (1), that there was a change in X, is established through a manipulation check. Manipulation checks will be discussed later in this chapter. Whether X´ is an appropriate exemplar of the theoretical independent variable (requirement (4)) will be discussed under Another Systematic Source of Variance: the Independent Variable.

Condition (2) (whether there was a change in Y) refers to the sensitivity of the measure of Y and to the statistical significance of the observed change in Y´. This requirement will not be discussed further in this chapter, nor will be conditions (5)(which concerns the construct validity of the criterion variable) or (8) (which regards the theoretical convergence of X and Y; however, see Baker and Churchill (1977)

as an example of a study in which attractiveness is both the independent variable and a component of the criterion variable. This is what Borg and Gall call "statistical contamination" (1979, p. 170)).

Requirement (7) regarding the generality of the statement $\underline{\Delta X} \rightarrow \underline{\Delta Y}$ will also be discussed later in this chapter.

Requirement (3), the priority of $\underline{\Delta X}$, is met through the manipulation of the independent variable in an experiment. Thus, the ambiguity regarding directionality in the statement, $\underline{\Delta X} \rightarrow \underline{\Delta Y}$, which would be present if X were measured (as in correlational studies) will be avoided. "In our laboratory experiment...we know what caused X--the experimenter did" (Carlsmith, Ellsworth and Aronson 1976, p. 34).

Alternatives to $\underline{\Delta X}$ as an explanation of $\underline{\Delta Y}$ (i.e., a violation of requirement (4) may occur in two ways. First, it may arise from variables which are covariate (confounded) with X within or across experiments. Confounding will be discussed under Systematic Sources of Variance in the Criterion Variable: Confounding. Second, it may arise from the covariance between Y and some third varible which is influenced by a $\underline{\Delta X}$ (see Birnbaum and Mellers 1979).

How Requirements for Valid Inference are violated:

At the simplest level $\underline{\Delta X}$ may be accomplished by having X occur at the zero level in one condition (control) and at a non-zero level in the other condition (treatment). Given no other influences on Y, the difference in Y for the two condition would then be considered as substantiating evidence for the hypothesis. "If X, then Y [treatment condition]...coupled with if not-X, then not-Y [control condition]... the two formulas together constitute Mill's joint method of agreement and difference, establishing the independent variable, X, as both the sufficient and the necessary condition of Y" (Boring 1969, p. 1).

Unfortunately, the paradigm cannot be as simple as that. First of all, Y will often have some initial value. Secondly, there will always be other variables besides X that differ between the treatment and control conditions. For example, the treatment and control conditions cannot exist simultaneously for the same person. If different people are used as subjects, then there will be personological variables (Bracht and Glass, 1968) that also differ between conditions. If the same person is subjected to the two conditions at different points in time (e.g, not-X followed by X), then aspects of the individual as well as external variables that change with the passage of time (i.e., maturation and history, Campbell and Stanley 1963) will differ between the two conditions.

These and other variables which differ between conditions may influence the criterion variable either singly or in concert with the independent variable. "...it is highly unlikely that any psychologically discriminable stimulation which we apply to an experimental subject would exert literally zero effect upon any aspect of his performance" (Meehl 1967, p. 109). Given that no two conditions can differ only in X and that any discriminable variable may influence the criterion variable: "It is impossible to design an experiment in which no variables except the independent variable affect the outcome..." (Carlsmith, Ellsworth and Aronson 1976, p. 13).

The presence of alternative explanations for ΔY is a problem that plagues experimental scientists. This problem is exacerbated by the faulty logic by which scientists draw causal conclusions. Mahoney (1976) states that there are two valid forms of implication which follow from the rules of propositional logic:

> The first form of implication is known as <u>modus ponens</u>, or confirmatory reasoning, and second is called <u>modus tollens</u>, or disconfirmatory reasoning. According to the rules of logic..., these are the <u>only</u> valid forms of conclusive inference. Note that they embody the crux of formal logic: a true premise must yield a true conclusion, and a false conclusion necessarily implies a false premise. Unfortunately, humans--and, as we shall soon see, scientists--are not inherently logical creatures. Their two most common errors are <u>affirming the consequent</u> (invalidly reasoning that a true consequent implies a true premise) and <u>denying the antecedent</u> (inferring that a false premise must imply a false consequent). Now then, let us look at the most typical strategy of theory evaluation in science. The scientist arrives at a theory or hypothesis (p) which predicts some experimental or observational phenomenon (q). He (implicitly or explicitly) reasons that "if my hypothesis is true, then I will observe such-and-such." When he subsequently makes these observations, he illogically concludes that they "verify" his hypothesis. That is, <u>the reasoning used in theory evaluation is often the illicit form of affirming the consequent</u>! ...In a strictly technical sense, disconfirmation is the <u>only</u> form of valid conclusive inference currently available to the scientist. This is because <u>successful predictions, no matter how numerous, have no necessary logical bearing on the truth status of the theory or hypothesis from which they were derived</u>. They may spawn <u>subjective</u> confidence in the sense that the hypothesis has repeatedly escaped disconfirmation, but until we develop other forms of valid scientific inference-- <u>it is only unsuccessful predictions which have conclusive logical implications</u> (pp. 138-139).

Successful predictions are not conclusive because $\underline{\Delta X}$ is only one of many variables that comprise the set of possible causes of $\underline{\Delta Y}$. Thus, the syllogistic reasoning followed by scientists leads to a conclusion that may be invalid. This type of invalidity has been labeled <u>internal invalidity</u> (Campbell 1957). "Drawing false positive or false negative conclusions about causal hypotheses is the essence of internal validity..." (Cook and Campbell 1976, p. 231). This invalidity is one aspect of what Kruglanski calls <u>treatment invalidity</u> (1975b, 376).

Most experiments use multiple subjects per condition and may utilize different rooms, times of day, experimenters, etc. As more variables are added to the experimental situation, some will vary systematically with the independent variable and some will vary

independently of the independent variable within and across conditions. The next two sections will consider first the systematic sources of variance in the criterion variable and then those sources independent of the independent variable.

Systematic Sources of Variance in the Criterion Variable: Confounding

Systematic extraneous variables take on different values for the different conditions in an experiment. Since the independent variable(s) also (by definition) take(s) on different values for the different conditions in an experiment, systematic extraneous variables co-vary with the independent variable. Such variables are said to be confounded with the independent variable (see Keppel 1973, p. 5; Carlsmith, Ellsworth, and Aronson 1976, p. 24). "Systematic confounding means that each level of the independent variable is accompanied in some consistent way with one of the levels of another variable" (Gottsdanker 1978, p. 53).

Different Types of Confounding:

Gottsdanker has enumerated several different types of confounding. One type is proceedural confounding. Proceedural confounding arises "...because the same subject cannot be given the different treatments simultaneously" (Gottsdanker 1978, p. 158). In this type of confounding, variables such as rooms, subjects, experimenter, or time of day covary with the independent variable.

Another broad category of confounding detailed by Gottsdanker is associative confounding. Associative confounding results from the experimenter's inability to manipulate just one thing in operationalizing an independent variable--that is, because the treatments are impure. Some associative confounds arise because of the intervention of

the experimenter--"...from the special manipulations needed to produce one of the experimental treatments" (Gottsdanker 1978, p. 166). Gottsdanker calls these artifactual confounds. They would not occur under normal circumstances.

A second type of associative confound is what Gottsdanker calls a natural confound. One type of natural confound involves the level of specificity at which the independent variable is labeled. The manipulation of an independent variable at a specific level unavoidably involves the manipulation of a more general variable (what Gottsdanker calls a wider variable, 1978, p. 172) that subsumes the intended variable. For example, watching the speaker present a message in an experimental setting also could be viewed as watching someone while the message was presented.

A second type of natural confound involves secondary variables. Rather than applying to a more general variable which subsumes the intended predictor variable, a confounding secondary variable involves another variable which naturally covaries with the independent variable (Gottsdanker 1978, pp. 173-174).

Another type of confounding refers to the presence of one level of an extraneous variable across all conditions within one experiment, but different levels of the variable across experiments which examine the same phenomena. Instead of a conditionwise confounding, this represents an experimentwise confounding. An example would be the use of one experimenter in all conditions within one experiment but a different experimenter in other experiments. The implications of this as well as other sorts of confounding will be discussed in the next section.

Problems of Inference Which Result from Confounding:

The presence of confounding in an experiment provides alternative explanations for any observed changes (or non-changes) in the criterion variable. If these alternative explanations are not discovered they can result in errors of inference (i.e., inferring that the independent variable caused the change in the criterion or failed to cause a change). Variables confounded with the independent variable and which could influence the criterion variable are called artifacts (the term, artifact, has also been used to denote the error of inference (Kruglanski 1975a, p. 103) and also the effect of the confounded variable (Rosenthal and Rosnow 1975, p. 220; Carlsmith, Ellsworth and Aronson 1976, p. 24). An artifact can either have a main effect on the criterion variable or may interact with the predictor variable. "...The term artifact refers to a factor producing a spurious relationship between variables, thus leading to invalid inference in a scientific inquiry. An artifact could denote a factor confounded with the independent variable in experimental research, threatening the validity of a conclusion that the latter was accountable for the observed effects (internal validity). Alternatively, an artifact might invalidate the inference that the results of a given research are generalizable to other, yet untested situations (external validity). To constitute an artifact, a variable needs to satisfy two conditions: (a) be capable (singularly or in interaction) of influencing the phenomenon under study and (b) be systematically associated with some of the experimental treatments or with the background conditions of an experiment (or experiments)" (Kruglanski 1973, pp. 348-349). Thus, an artifact may lead to an error of inference because it has a main effect on the

criterion, because it interacts with the independent variable within the experiment, or because it interacts with the independent variable between experiments.

Inferential errors in which the artifactual variable has a main effect on the criterion variable.

When an artifact has a main effect on the criterion variable it may lead to two different types of inferential errors. If the independent variable failed to influence the criterion variable, the effect of the artifact may be misinterpreted as being that of the independent variable. If the independent variable has an effect on the criterion, then whether or not an error of inference results depends on the directionality and magnitude of the effect of the independent variable and of the artifact. If they are in the same direction, the supposed singular effect of the independent variable will simply be exagerated, and no error of inference will result. However, the extraneous variable may account for most of the change in the criterion, with that portion accounted for by the independent variable being non-significant. In such a case the effect could erroneously be attributed to the independent variable. If the directionality of the effects of the independent variable and of the artifact are in the opposite direction, then a true effect may fail to be detected.

Inferential errors in which the artifactual variable and the independent variable interact within an experiment:

The differences in the criterion variable between conditions (or lack of differences) may be caused by interactions of the independent variable with artifacts of the study. If the level of the artifactual variable influences the relationship between the independent variable

and the criterion, then the effect of level a_1 of the independent variable <u>in the presence of</u> level b_1 of the artifact may differ from the effect of level a_2 of the independent variable <u>in the presence of</u> level b_2 of the artifact, even though the effects of levels a_1 and a_2 may be the same if accompanied by the same level of the extraneous variable. In a similar fashion, real effects of the independent variable may be obscurred by interactions with artifacts.

Kruglanski (1975b) suggests that in some cases the <u>combinations</u> of a level of the independent variable with a level of an artifact may <u>define</u> a relevant variable which merits study in its own right. In fact, complex independent variables often involve intentional "confounding" of variables (Underwood and Shaughnessy 1975, p. 31).

<u>Inferential errors in which extraneous variables are confounded with experiments</u>:

Even if an extraneous variable is constant across conditions, differences across <u>experiments</u> may lead to inferential errors (see Cronback 1975, p. 123; Carlsmith, Ellsworth and Aronson 1976, p. 88; and Weber and Cook 1972, p. 291). The differential effect of different levels of the independent variable may not be observed in an experiment in which the extraneous variable is held at one level and may be observed in another experiment in which the extraneous variable is held at another level. Again this may be the result of an interaction, or the combination of the independent and the extraneous variables may define another relevant variable (Kruglanski 1975b, p. 376).

<u>Another Systematic Source of Variance: The Independent Variable</u>

Other than extraneous variables which are confounded with the independent variable, the other source of systematic variance in the

criterion variable of an experiment is the independent variable. In an experiment, the independent variable is manipulated, that is, some operation is performed to effect its occurrence. The operation(s) comprise(s) what Carlsmith, Ellsworth, and Aronson call the "empirical realization" of the independent variable (1976, p. 11).

Choosing Appropriate Operations:

The approach to choosing the operations to effect the independent variable depends on the nature of the research. Problematic research often involves complex variables (as in advertising copy testing) in which it is not clear what the component (usually confounded) variables are (Underwood and Shaughnessy 1975, p. 31).

Other times research may be concept-oriented and "programatic" without there being a clear notion of the theoretical significance of the concepts. "The interest in recent years has been in building a firm body of knowledge about a given concept. Whether a variable is suggested by one theory or another is less important than its ease of operational definition and relevance to the problem...The specific researchable question may take over to the exclusion of any theoretical concern and the results may not contribute significantly to the understanding and prediction of human behavior" (Katz 1972, p. 557).

For theory-oriented research, careful consideration should be given to the nature of the independent variable as suggested by the relevant theory. "The choice of operations should depend on the result of a conceptual analysis of the essential features of a construct... A precise explication of constructs is vital for high construct validity since it permits tailoring the manipulations and measures to whichever definitions emerge from the explication" (Cook and Campbell 1976, p. 241-242).

Translating definitions into operations is a very difficult task. From the point of view of the cumulative development of the understanding of behavioral phenomenen this process is complicated by the multiple interpretations which those difinitions may have. "To a logician, X is X, but to a scientist X is a variable, identified in words which may easily mean one thing at one time and another at another, or different things in different laboratories or at different periods of history" (Boring 1969, p. 6)

The relationship between the independent variable and its exemplar (operation) in an experiment depends on whether the independent variable is defined as a stimulus or as a response on the part of the subjects. If the independent variable is defined as a stimulus, it may be one which is very concrete or it may be relatively abstract.

If it is very concrete (such as the number of presentations of an advertisement) then the "...concept and the event...[may be] so closely linked as to be almost identical" (Carlsmith, Ellsworth and Aronson 1976, p. 134).

If the independent variable is a relatively more abstract stimulus, then the operation used to effect it may be an "...event taken as representative of a broader class of events which are included in the concept..." (Carlsmith, Ellsworth and Aronson 1976, p. 134).

If the independent variable is defined as a response, the operation is not even an instance of the independent variable, but is intended to induce it as a response by the subjects (see Carlsmith, Ellsworth and Aronson 1976, p. 62).

When the independent variable is a response it is called a conceptual variable (Carlsmith, Ellsworth and Aronson 1976, p. 62) and

sometimes an intervening variable (see Underwood and Shaughnessy 1975, p. 133). However, MacCorquodale and Meehl (1948) argued that the term, intervening variable, should be reserved for variables which have no purported existence beyond the empirical terms by which they are defined. Instead MacCorquodale and Meehl suggested that variables such as response-type independent variables be called hypothetical constructs which are "...not wholly reducible to empirical terms; they refer to processes or entities that are not directly observed (although they need not be in principle unobservable); the mathematical expression of them cannot be formed simply by a suitable grouping of terms in a direct empirical equation; and the truth of the empirical laws involved is a necessary but not sufficient condition for the truth of these conceptions..." (1948, p. 104).

Establishing the dimensionality of the independent variable:

The choice of specific operations must be guided by the conceptual analysis of the independent variable which was discussed earlier. A review of the relevant literature, as well as dictionary definitions and commonsense notions, will suggest the dimension(s) which comprise(s) the variable. For example, in the present study it was decided to utilize Kelman's (1953) definition of credibility which specifies the dual dimensionality of trustworthiness and expertise. This is an example of a complex independent variable which consists of the intentional "confounding" of the constituent variables, e.g., high credibility would consist of high trustworthiness and high expertise while low credibility would consist of low trustworthiness and low expertise. (See Underwood and Shaugnessy 1975, p. 32). The other two combinations of the constituent variables were not considered.

After the representative dimensionality of the independent variable is chosen, the intended levels must be established. Here the researcher should keep in mind that "...the strength to which different variables can be produced [in an experiment] is extremely weak compared to the strength with with these variables exist and operate in real-life situations" (Festinger 1953, p. 141). This affects the power of an experiment, i.e., the probability of rejecting a false null hypothesis (Keppel 1973, p. 528). In other words, if the manipulation of the independent variable is too weak, the experiment may not provide a test of the proposed relationship that is sensitive enough to detect true effects of $\underline{\Delta X}$ on $\underline{\Delta Y}$. Sensitivity of an experiment can be increased by having stronger manipulations of the independent variable. This suggests widely divergent levels between conditions along the selected dimensions.

The use of widely divergent values of the selected dimension has another benefit. "In general, the more limited the range of values as compared with the range of universe values the greater the potential loss of information for the relevant variable" (Fromkin and Streufert 1976, p. 430). Additional information regarding the relationship between the predictor and criterion variables across the range of possible values of the predictor variable may be gained through the use of intermediate as well as extreme values of the predictor. This is expecially important if there is reason to believe that the relationship between the predictor and criterion variables is non-linear. Three levels of discrepancy were used in the present study for that reason.

There are two restrictions to consider in the use of extreme values of the predictor variable. One is an ethical consideration. Subjects

should not be exposed to levels of a variable that could cause physical or psychological harm (see Cozby 1977, p. 78 and Carlsmith, Ellsworth and Aronson 1976, p. 83). The second consideration deals with whether the manipulation is realistic. "...The strongest possible manipulation may involve a situation that rarely if ever occurs in the real world..." (Cozby 1977, p. 78).

Most behavioral studies attempt to establish whether there is any relationship between variables and the general nature of the relationship, rather than attempting to calibrate, in a metric sense, the amount of change in the criterion variable that might be expected in a natural setting for a specific change in the predictor. In other words, few behavioral studies attempt to estimate parameters in the population, but rather involve inexact hypotheses (Payne 1977) and general- rather than aggregate-type propositions (Bakan 1955). Even if a study were to attempt to predict the amount of change in Y in a natural setting, if Festinger is right about the differential effect of an operation in an experimental and a naturalistic setting, then such predictions would be extremely difficult to make. However, across experiments the relationship between the predictor and criterion variables may only obtain when extreme values of the predictor variable are used, but then, this the reason for using extreme values.

The duration of an operation is another important consideration. It may be either too short or too long. As with many other aspects of experimentation, operations which are too short are most troublesome when significant changes in the criterion are not obtained. The possibility then remains that the operation was too short in duration to have any observable impact on the subjects. Operations which are too

long can also be troublesome. Extended operations provide more opportunity for the intrusion of extraneous variables that may cause either confounding or an increase in error variance. (Error variance is discussed in the later section: The Impact of Random Effects.)

Modes of manipulating the predictor variable:

The predictor (independent) variable may be manipulated verbally, by events to which subjects are exposed, or through changes in environmental factors. Verbal manipulations may be a part of the instructions (Carlsmith, Ellsworth and Aronson 1976, p. 147) given to the subjects or may be a part of a verbal text which is presented to them either orally or in written form. Verbal manipulations may comprise the exemplar of the independent variable or may be intended to induce the independent variable as a response.

Examples of stimulus-type, verbal manipulations would be messages in which the content or organization is varied or instructions to the subjects which include elements intended as exemplars of the independent variable. In the present study, discrepancy was manipulated through the content of a printed message.

Verbal materials are also used to effect responses as the independent variables of interest. In the present study verbal description was intended to effect the response of perceived credibility. False feedback (Carlsmith, Ellsworth and Aronson 1976, p. 149) regarding a subject's performance on a task is sometimes used to engender a temporary change in a subject's perception of his abilities or qualities (Underwood and Shaughnessy 1975, p. 28).

Events can also be used as operations intended to either be or bring about the exemplar of the independent variable. Events, however,

are more ambiguous than verbal material and their interpretation may be expected to vary more across subjects than those of verbal materials (Carlsmith, Ellsworth and Aronson 1976, p. 147).

An event-oriented variable may involve the subject as a participant or may occur independently of him. A participatory, event-oriented operation may involve a task to be performed by the subject (Underwood and Shaughnessy 1975, p. 27). These tasks may themselves represent the independent variable or may be intended to induce a response representative of the independent variable. Various distraction tasks, for example, are intended to manipulate the allocation of attention--a response (see chapter 3).

Examples of non-participatory events are various staged "accidents" (Carlsmith, Ellsworth and Aronson 1976, p. 152) and unexpected distractors such as the sound track in the Silverman and Regula study (1968).

Environmental factors such as temperature, humidity, lighting or time may also be manipulated to effect the independent variable (Underwood and Shaughnessy 1975, p. 28).

Problems of the meaning of the independent variable:

As was discussed at the beginning of this chapter two aspects of ΔX must be established before the statement $\underline{\Delta X} \rightarrow \underline{\Delta Y}$ can be justified. The first is : has there actually been a change in $\underline{X}$. The second is: is the operation an exemplar of $\underline{X}$ and not some other variable. The first issue will be dealt with later in this chapter in the discussion of manipulation checks. The second issue concerns two sources of construct invalidity. "They all have to do either with the operations failing to incorporate all the dimensions of the construct, which we might call

'construct underrepresentation' or with the operations containing dimensions that are irrelevant to the target constructs, which we might call 'surplus irrelavancies' (Cook and Cambell 1976, p. 241; see also Festinger 1953, p. 139).

In the present study, the credibility of the message source was intended to include both the perceived expertise and the perceived trustworthiness of the source. However, a manipulation check revealed that only perceived expertise had been affected by the manipulation (a description of the source). This is an example of construct under-representation.

Surplus irrelevancies consist of confounded variables. These artifacts may be either stimulus oriented or response oriented. The confounding may be just within or may be across experiments.

Surplus irrelevancies may be the result of experimental proceedures (see the earlier discussion of procedural confounds (Gottsdanker 1978). "By the time the subject is ready for the experimental task he may have been subject to such a range of influences, to establish an experimental condition, that it is not clear what the experimental condition really is" (Katz 1972, p. 556).

Surplus irrelevancies could also derive from trying to induce a response representative of an independent variable that often covaries with another conceptually distinct variable. Subjects may perceive both variables even though only one was intended. For example, it may be difficult to manipulate the perceived attractiveness of a source without also manipulating perceived credibility.

Random Effects

The Impact of Random Effects:

Variables which cause random effects on the criterion variable differ from those causing systematic effects in that they vary within conditions rather than between conditions. That is, they cause variation in the criterion scores of individuals within a condition about the mean criterion score for that condition. However, they do not affect the mean and, therefore, do not contribute to differences between means. This variation in the criterion that results from random effects contributes to the error variance of the experiment as opposed to bias which results from systematic (confounded) variables. "Bias refers to factors which systematically affect the value of means..." (Cook and Campbell 1976, p. 231).

Since random effects do not affect the means of the criterion variable across conditions, they cannot cause differences which could mistakenly be attributed to the independent variable. Nor could random effects act in a direction opposite to that of the predictor variable and cancel any true effects of the independent variable on the criterion means. However, error variance may obscure true effects of the independent variable because of the way the statistics used to test hypotheses are derived.

Many of the statistics used to test the effects of the independent variable involve a ratio of "differences among means" to "differences among subjects treated alike" (e.g., random effects; see Keppel 1973, p. 33). Generally, as the size of the ratio increases, the more confident the experimenter is that he (she) can reject the null hypothesis of no differences among means with a small likelihood of making an error.

However, given the same differences among means, an increase in error variance will decrease the magnitude of the ratio and decrease the confidence that the experimenter would have when rejecting the null hypothesis that it was not true.

As a result, increases in error variance increase the chance that a false null hypothesis will be accepted, thereby obscurring the true effects of the independent variable. This is a decrease in the power of the experiment.

Sources of Random Effects:

Random effects have many sources. For example, random assignment of subjects ensures that the random distribution of personological variables (Bracht and Glass 1968) across conditions. However, the greater the variability of subjects "...with respect to variables that affect the dependent variable," the greater the error variance (Cook and Campbell 1976, p. 233).

Another source is random variations in the conduct of the experimenter across subjects in single-subject presentations or across groups for group presentations (see Friedman 1967; also Keppel 1973, p. 31).

If the experimental stimulus is familiar to the subjects "...our stimulus situation is certain to interact...with individual differences among the subjects which arise from subjects' differing past experiences in similar situations...the subjects' interpretation of these events could be influenced by prior experiences beyond our control, and thus the 'same'situation would seem very different to different subjects" (Carlsmith, Ellsworth and Aronson 1976, p. 55). These varying interpretations would be a third source of random effects.

A fourth source of error vairance is the initial value (i.e., pre-treatment) of the criterion variable for experimental subjects. The

post-treatment measurement of the criterion variable will almost certainly reflect the starting value for subjects on that variable. If subjects are randomly assigned to conditions, then pre-treatment values of the criterion should not be s source of systematic variance in the post-treatment values of the criterion. Both the main effects and interactive effects with the treatment should contribute to error variance.

A fifth source of error variance (although non-random) is extraneous variables that have been balanced across conditions (e.g., "unanalyzed control factors" Keppel 1973, p. 523). Their balancing eliminates bias, but increases non-systematic error.

Generality of Research Findings

Generalization of research findings concerns interactions. The question of whether the observed relationship between a predictor variable and the criterion can be generalized may be restated as: Will the same relationship between the predictor and criterion be observed in situations in which extraneous variables present in the experimental situation take on values different from than those which the Y possessed in the initial experiment?

One type of generalization is statistical generalization which concerns generalizing from the sample which was tested to the hypothetical (or real) population from which the sample was taken. This generalization generally involves hypothesis testing or parameter estimation for the population. This generalization is based on the assumption of random selection of the sample from the population (a condition rarely met) and applies to the response of the population performing the same activities under the same conditions as the tested sample.

This sort of generalization is called statistical inference. Statistical inference is "...a way of making inferences about the population-aggregate from the study of the sample-aggregate" (Bakan 1955, p. 212). This type of generalization deals with aggregate-type propositions, but most behavioral studies examine the plausibility of general-type propositions. "A general-type proposition asserts something which is presumably true of each and every member of a designable class. An aggregate-type proposition asserts something which is presumably true of the class considered as an aggregate" (Bakan 1955, p. 211).

Most behavioral experiments do not legitimately allow these generalizations. "Almost invariably, our subjects are selected out of convenience, rather than at random. The failure to sample randomly from a known population means that we are not justified statistically in extending our results beyond the bounds of the experiment itself" (Keppel 1973, p. 28). This is one aspect of what Bracht and Glass refer to as population validity (1968, p. 438).

The type of propositions which are generally at issue in a behavioral experiment are general-type propositions. "General-type propositions are...critically testable, since they are jeopardized by each new instance. If a general-type proposition fails to be confirmed by the observation of a member of the class to which the proposition presumably applies, then either the proposition must be rejected, or the class must be more closely delimited. The 'next' case, for the aggregate-type proposition, simply increases the 'power' of the test, and the likelihood of the empirical proposition which is under consideration, if the study is properly conducted with respect to randomness, etc." (Bakan 1955, p. 212).

It is the non-statistical generalization of general-type propositions which is most often sought in behavioral studies. Rather than generalizing to the population from which the sample was drawn, the desire is to know to what <u>other populations</u> the results of the experiment may be projected. If the results are population-specific, then, of course, this represents an interaction and limits the generalizability of the findings.

It is in this way that any interaction between the independent variable and extraneous variables limits the generalizability of experimental findings. The presence of interactions restricts statements regarding the relationship between the predictor and criterion variables to specific levels of the extraneous variables which interact with the predictor variable. Given the infinite number of extraneous variables which could interact with the independent variable, the problem of extending the experimental results to situations different from those in the experiment is a serious one (see Bracht and Glass 1968, p. 455). "By knowledge we mean, in part useable reidentifiable sameness in settings that are not identical. If the highest order interactions with the specifics of space, time, and attributes are always significant, then no generalization is possible, and hence no knowledge and no science" (Campbell 1969, p. 359).

There are several sources of interactions of particular interest besides population specific (i.e., personological) variables which may interact with the independent variable and thereby limit the generality of the findings. Personological variables comprise the second aspect of population validity outlined by Bracht and Glass 1968, p. 438. Recall that their first aspect deals with whether the sample is <u>representative</u> of the population.

One source of interactions is the extraneous variables which comprise the "surplus meaning" contained in the independent variable. The relationship between the independent and dependent variables may only occur for certain operationalizations, i.e., in the presence of certain levels of extraneous variables inherent in those operationalizations. As Kruglanski has pointed out, however, the combination of the independent and "extraneous" variables may jointly operationalize another variable which has theoretical relevance. "It will be noted that within theoretical research the separation of contexts [i.e., restrictions on generalizability] may be regarded as a practical limitation resulting from insufficient knowledge. Within such research the relentless quest is for maximally general laws, so context-separation...exists only as a temporary (even if sometimes protracted) challenge to be overcome in the future by an insightful redefinition of the critical variables" (Kruglanski 1975b, footnote p. 377).

A second source of interactions is the factors comprising setting of the experiment, i.e., what Weber and Cook call "procedure-correlated artifacts" (1972, p. 291). These include the presence or absence of other subjects, the physical surroundings, the "atmosphere," lighting, etc. "Under conditions of diminished novelty, the superiority of the experimental treatment may dissappear" (Bracht and Glass 1968, p. 459).

The experimenter is a third extraneous variable which may interact with the treatment and limit generalizability. Kintz et al. (1965) list many experimenter variables which influence subject responses to the independent variable. Therefore, findings may not be consistent across experimenters either within or across experiments.

A fourth source of interactions and, thus, a restriction on the generality of experimental findings is "the times.... Generalizations decay. At one time a conclusion describes the existing situation well, at a later time it accounts for rather little variance, and ultimately it is valid only as history" (Cronbach 1975, p. 122-123)

The preceding discussion has illustrated that valid conclusions regarding the proposition, $\underline{\Delta}X \rightarrow \underline{\Delta}Y$ may be threatened by confounded extraneous variables, by interactions between the independent variable and extraneous variables both within and across experiments, and by extraneous variables that vary randomly within conditions. To increase the confidence the experimenter has in his or her conclusions, the extraneous variables which can influence the criterion variable directly or can influence the relationship between the predictor and criterion variables must be controlled.

Control

Control in an experiment means the control of variation. Four types of control can influence the validity of conclusions regarding the experimental hypothesis: (1) control of confounding, (2) control of variables which may interact with the independent variable, (3) control error variance, and (4) control of variation in the independent variable.

Control of Confounding:

The basic problem is controlling the confounding of an extraneous variable and the independent variable is to eliminate the covariance between the two variables. However, there is an infinite number of extraneous variables which may be confounded with the independent variable. First, then, a decision must be made as to *which* extraneous

variables to control. "...It is not failure-to-control in general that bothers us, but only those failures of control which permit <u>truly plausible</u> rival hypotheses, laws with a degree of scientific establishment comparable to or exceeding that of the law our experiment is designed to test" (Campbell 1969, p. 356).

Campbell's suggestion may be somewhat optimistic, as it presumes knowledge regarding which variables may influence the criterion or the relationship between the predictor and the criterion. In some cases "...the confounding variable is one on which there has been no research; that is, it is not known whether the variable does or does not have an influence on the behavior being measured in the experiment. All we know is that its magnitude is correlated with the independent variable of interest. We tend to take a fairly hard line in these situations and conclude that the results are ambiguous as to their cause, and therefore they should not become a part of our literature" (Underwood and Shaughnessy 1975, p. 323). This attitude presumes an awareness of the variable and of its covariation with the independent variable.

Still, any method of control other than randomization requires that choices be made as to which extraneous variables are the "'most likely' influences..." on the criterion variable (Mahoney 1978, p. 665). Researchers must rely on theory, prior empirical studies, and on hunches based on knowledge of the subject matter to provide clues as to which extraneous variables may invalidate their conclusions if left uncontrolled. "One value of a theory is that it tells us whether the differences which do exist between...events are important and if so, in what ways" (Schlenker 1974, p.8). In addition, an "...awareness of other variables that may affect the behavior under observation may allow

the researcher to measure these variables and analyze the data so as to determine whether any of them were confounded with the primary variable of interest" (Carlsmith, Ellsworth and Aronson 1976, pp. 50-51).

The use of randomization to avoid confounding:

Randomization can simultaneously control many sources of bias. For example, subject variables such as past experiences, income, intelligence, and group membership may all be controlled by randomly assigning subjects to conditions. Randomization removes all of these subject variables (known and unknown) as systematic sources of variance in the criterion and makes them instead random sources of variance.

Randomization is so useful in controlling subject variables and in increasing the internal validity of an experiment that it is considered an essential element of a behavioral experiment (Carlsmith, Ellsworth and Aronson 1976, p. 15). Randomization can also be used to control for other extraneous variables by assigning individual subjects randomly, not only to conditions, but also to experimenters, rooms, times of day, etc. within conditions.

The negative aspect of randomization for variables other than subject variables is that it increases the error variance and, thereby, reduces the power of the experiment to detect effects of the independent variable (Keppel 1973, p. 314).

Avoiding confounding by holding extraneous variables constant across conditions:

A second way of avoiding confounding is to hold an extraneous variable constant at one level across conditions. (A special case is one in which the extraneous variable is held constant at the zero level across conditions.) This approach may mean adding a non-zero level of

the extraneous variable to a control condition or reducing the extraneous variable to zero in the treatment condition. Campbell calls the result either the "expanded content control group" or the "reduced content experimental group" (1969, p. 359). This procedure eliminates the possibility of a main effect of the extraneous variable on the criterion. However, the possibility remains of an interaction between the extraneous variable and the predictor variable leading to different effects _across experiments_ in which the extraneous variable takes on different values. As a result, holding an extaneous varaible constant increases the internal validity of the experiment but may decrease the generality of the findings, i.e., external and ecological validity (See Carlsmith, Ellsworth and Aronson 1976, p. 17).

Avoiding confounding through manipulation of extraneous variables:

Balancing. Another way of avoiding confounding is to artificially provide for an extraneous variable to vary in the same manner in all conditions. That is, it should take on the same values in each condition. McGuigan (1978, p. 155) calls this approach _balancing_ to distinguish it from _counterbalancing_, a term which he reserves for the alternation of order of presentation to eliminate order effects. Baker and Churchill (1977) is an example of a study in which the order of presentation is confounded with treatments). However, others call the approach _counterbalancing_ (e.g., Keppel 1973, p. 314; Fromkin and Streufert 1976, p. 432).

The result of balancing is what Kaplan refers to as _cancellation of error_. "...Errors are made first in one direction, then in the reverse direction,...so that the same error is made in both the test group and the control group" (1964, p. 156).

However, this error within both conditions, while not a source of bias, increases the error variance and reduces the power of the experiment.

In the present study all subjects had already been exposed to a prior treatment (with which this study was not concerned). To avoid confounding of prior treatment conditions with conditions of the present study, the prior conditions were balanced across conditions of the present study.

<u>The extraneous variable as an independent variable</u>. A method of eliminating bias while not contributing to the error variance is to include an extraneous variable in the experimental design and analysis as an independent variable, i.e., as a "control variable" (Keppel 1973, p. 523). When <u>balancing</u> is used the <u>effect</u> of the extraneous variable is not measured. However, when the extraneous variable is included in the analysis as an independent variable, its main effect on the criterion variable may be assessed, as can any potential interaction of the extraneous variable with the independent variable. This is not to suggest that the analysis is performed because the researcher is interested in the effects of the control variable. The primary purpose of the approach is to reduce the error variance and to reveal any interactions which may limit the generality of the effects of the independent variable (see Carlsmith, Ellsworth and Aronson 1976, p. 140 and Keppel 1973, p. 316). If this method is used instead of randomization, the result will be a smaller error variance and increased power of the experiment, which is a distinct advantage of this approach over balancing (in which the effects of the extraneous variable are not analyzed; see Keppel 1973, p. 522 for a discussion of "unanalyzed control factors").

"The only 'loss' is a reduction in power due to the fact that the control factors remove degrees of freedom from the within-group error term which would have been present if the sources involving the control factors had not been removed. In most cases, however, this loss of power is negligible" (Keppel 1973, p. 318).

Measuring the extraneous variable. If the variation in an extraneous variable cannot be delimited or randomized by the experimenter, its impact on the criterion variable can still be isolated by measuring the independent variable. Thus, even though the extraneous variable could not be included in the design it can be included in the analysis (see Kaplan 1964, p. 156). Its covariation with the criterion variable can be assessed and removed from the error variance and any interactive effects with the independent variable can be revealed. This increases the power of the experiment and indicates whether the effects of the independent variable may be generalized across the levels of the extraneous variable (Carlsmith, Ellsworth and Aronson 1976, p. 140).

Control of Interactions of Extraneous Variables with the Independent Variable:

Replication across experiments:

As discussed earlier, holding an extraneous variable constant within an experiment eliminates bias, but does not remove interactive effects which may limit the generality to settings in which the extraneous variable takes on other values. In other words, the observed effect of the independent variable may only occur when the extraneous variable has the value represented in the experiment.

To eliminate the possibility of experiment-specific results, experiments should be replicated with the extraneous variables taking on different values than they did in the initial experiment.

"Exact replication of all experimental conditions is not desirable in ...[behavioral] research because then the theory-relevant aspects of the treatment cannot be separated from the extraneous variables" (Bracht and Glass 1968, p. 453). Replications can best serve the goal of establishing the generality of an effect through the use of "...a heterogeneous sample of subjects, a wide variety of problems, a wide variety of settings,...[and] a wide variety of response measures" (Carlsmith, Ellsworth and Aronson 1976, p. 89).

Replication within experiments:

The problem of generality also applies to the particular operationalization of the independent variable used in the experiment. "Since single operations both underrepresent constructs and contain irrelevancies, construct validity will be lower in single exemplar research than in research where each construct is multiply operationalized in order to triangulate on the referent" (Cook and Campbell 1976, p. 242). Practical restrictions, such as cost, the availability of subjects, etc., may limit the use of multiple operationalizations within an experiment.

If the effects in the initial experiment attained significance and were published, there may be limited publishing possibilities for a replication of the experiment. As a result, few researchers may be motivated to extend the generality of the initial findings through replications. This is unfortunate, since "the most important syntheses and unifications in science depend on the introduction of theoretical concepts through convergent avenues of specification, coming from different, often at first very heterogeneous areas of evidence" (Feigl 1959, p. 127).

In summary, generality may be increased by increasing heterogeneity within a single experiment (Keppel 1973, p. 315 and Carlsmith, Ellsworth and Aronson 1976, p. 89), but with an accompanying decrease in the power of the experiment. The alternative is to introduce heterogeneity across experiments through replication.

Control of Sources of Error Variance:

As mentioned previously, the larger the error variance, the lesser will be the power of an experiment, i.e., the less sensitive it will be in detecting effects of the independent variable. Therefore, one goal of a researcher should be to minimize the error variance in his experiment.

Error variance results from the variation, within conditions, of variables which influence the criterion variable. Consequently, error variance will increase as the variation of these extraneous variables increases or as more variables are allowed to vary within conditions. Error variance may be reduced by reducing the variation of these variables within conditions or by reducing the number of variables allowed to vary within conditions.

Some of the methods used to control confounding increase error variance while others decrease it. Both randomization and balancing of extraneous variables increase error variance. Holding an extraneous variable constant across conditions, introducing the extraneous variable as an independent (control) variable, and measurement of the extraneous variable (followed by covariance analysis) all serve to reduce error variance and increase the power of the experiment.

"There are three major sources of error variance: random variation in the actual treatments, the presence of unanalyzed [balanced] control

factors, and individual differences ('permanent' or 'temporary' factors affecting a subject's performance during the course of the experiment)" (Keppel 1973, p. 522).

The control of error variance arising from treatment variability involves reducing the variability within conditions of extraneous variables associated with the treatment. This means that lighting, background noise, the presence of others, the behavior of the experimenter, etc. should all be standardized across subjects. This could be accomplished by running all conditions, simultaneously, in the same room by one experimenter. However, practical considerations may make this impossible and special efforts must then be taken to ensure uniformity of conditions across subjects. "...Well-trained experimenters, and special testing rooms help to accomplish this goal" (Keppel 1973, p. 523).

Another source of variation in treatments is the interpretation of the treatments by the subjects. Response-oriented variables (hypothetical constructs) may generate different responses across subjects. The stimuli should be made as unambiguous as possible to reduce this effect.

However, Carlsmith, Ellsworth and Aronson (1976) suggest that in some cases a flexible treatment may be necessary when the independent variable is a hypothetical construct. They suggest that what is important in such experiments is response standardization rather than stimulus standardization. "Our recommendation is that the experimenter recognize that variability exists in the subject's understanding of the instructions and that on some occasions a standardized manipulation may not fit" (Carlsmith, Ellsworth and Aronson 1976, p. 162). This would

require some means of monitoring the independent variable as the manipulation progresses, a situation that is not feasible for many hypothetical constructs. However, excess variation in the interpretation of experimental stimuli may be detected through the use of a manipulation check during a pilot study which would allow adjustment of the stimuli to achieve greater uniformity in response prior to running the experiment.

Error variance which arises from subjects' pre-treatment values on the criterion may be controlled by pre-treatment _measurement_ of the criterion variable and the use of this pre-treatment value as a covariate in the analysis. This not only reduces the error variance but also will detect any interactive effects between the treatment and the pre-treatment criterion values. However, pre-treatment measurement of the criterion variable may have adverse cueing and facilitating effects on the criterion variable (see the later discussion of questionnaire effects, p. 69).

The increase in error variance which derives from unanalyzed (balanced) control factors may be controlled by making the extraneous variables into independent variables in the design _and_ analysis. This also allows any interactions of the control variables with the predictor variables to be assessed.

Control of the Variation in the Independent Variable:

To help avoid errors of inference regarding the relationship being examined in an experiment there are three desirata of an operationalization of an independent variable: (1) the operationalization should contain a _minimum of confounded extraneous variables_, (2) it should effect widely _disparate values_ of the independent variable, (3) the

independent variable should be effected <u>uniformly</u> within conditions. Confounding may lead to falsely attributing changes in the criterion variable to changes in the predictor variable. The use of levels of the independent variable that are too close together and the variation in the independent variable within conditions may both lead to falsely concluding that the independent variable does not affect the criterion variable.

It has already been discussed how each of these deserata may be influenced. This section will discuss how the success of the experimenter's efforts may be assessed.

<u>Assessing variation of the independent variable: manipulation checks</u>:

All three desired qualities of the operationalization of the independent variable may be assessed through the measurement of the independent variable. Such measurement is called a manipulation check (Festinger 1953, p. 146). It may consist only of measurement of the independent variable but may also include the measurement of extraneous variables which are likely confounds of the independent variable, thus helping to ensure the relative purity of the manipulation (Cook and Campbell 1976, p. 239).

<u>Timing of the manipulation check</u>. The manipulation check may be performed in the experimental sessions proper or it may be performed in a pilot study.

Whenever possible the manipulation check should be performed in a pilot study. There are several advantages to doing this. The foremost is that "...if the check shows that your manipulation was not effective, you have saved the expense of running the actual experiment" (Cozby

1977, p. 85). A second advantage is the flexibility that may be exercised in running a pilot study. It is not necessary to run through the entire experiment before taking measures of the independent or extraneous variables. "...It is a good idea to interview the pilot subject right after the treatment, without continuing the whole experiment. If one waits until the experiment is over and then attempts to question the subject about the dependent variables as well, the subject may find it difficult or impossible to describe the effects of the experimental treatment" (Carlsmith, Ellsworth and Aronson 1976, p. 163).

Another advantage to performing the manipulation check in a pilot test is that it may avoid consistency effects that could affect the measure if the manipulation check were collected during the experiment. For example, what subject would want to state that a source was non-credible when he or she had just indicated agreement with that source. However, not measuring the dependent variable does not guarantee that consistency effects will be absent from the measure of the independent variable. If the experiment is run in its entirety, subjects' evaluations of the credibility of a source may still be colored by whether they were <u>persuaded</u> by that source whether or not they must indicate such persuasion. This is another argument for stopping the pilot study after the manipulation to assess its impact. For example, the credibility of a source could be measured after its manipulation, but <u>prior to</u> any message from that source. That way subjects' agreement with the source could not affect their evaluation of credibility.

Pilot tests are not without their costs. If there are several independent variables, with several levels each, then the pilot study

may require a large number of subjects to assess the effectiveness of each manipulation.

Another cost is the time that must be devoted to the pilot study. However, a pilot study serves other important functions that justify its being conducted. The pilot study gives experimenters a chance to practice and, thereby, reduce variability in their conduct. It also provides an opportunity to determine if the instructions are clear, that the tasks may be performed, that the questionnaires are unambiguous and may be completed, and that all may be completed within the desired time frame. All of this may be accomplished with few subjects, --perhaps far fewer than may be necessary to assess the success of multiple manipulations.

A shortcoming of the pilot study as a measure of the success of the independent variable manipulations is that success in the pilot study does not guarantee success in the experiment proper (Fromkin and Streufert 1976, p. 447). The results of pilot studies are most meaningful when they indicate the <u>failure</u> of the manipulations. Then the manipulations may be revised so that the research proposition may be adequately tested in the experiment.

Manipulation checks taken within the experiment proper reveal whether the manipulation appears to have been successful for that specific sample in that particular setting. However, there are several problems associated with performing manipulation checks within an experiment. One revolves around the timing of the measures. Measures of the independent variable taken early in the experiment are less likely to be contaminated by consistency effects, but are more likely to contaminate the criterion measure. "A more obvious situation in which a

question can act as an artifact is an experiment that employs deception; in this context, asking subjects to report on an internal process is very likely to make them aware of the purpose of the experiment" (Carlsmith, Ellsworth and Aronson 1976, p. 73).

Because of the seriousness of potentially influencing the criterion measure, manipulation check measures taken in an experiment should follow the criterion measure. However, because of possible consistency effects the results of the manipulation check will be somewhat ambiguous. Again they will be most meaningful if they indicate that the manipulations were unsuccessful. This would help avoid falsely concluding that the predictor variable does not influence the criterion variable. "Laws and hypotheses are formulated, either explicity or implicitly, in a logical implicative, if-them form; if particular antecedent conditions are realized then particular consequences should follow. If the antecedent conditions are not realized, either because of improper methodology, a change in the values of parameters, the intrusion of unforeseen events, etc., then there is no reason to expect the consequents to follow" (Schlenker 1974, p. 7). If the manipulation check indicates that the manipulations were unsuccessful, then changes could be effected for subsequent experiments.

However, it is also possible to falsely conclude that the manipulation was not effective. "Following an experiment, it is not at all uncommon for subjects to deny any feelings of the kind the experimenter expected to arouse, while at the same time their behavior throughout the experiment was just what would be anticipated if they had experienced precisely those feelings" (Carlsmith, Ellsworth and Aronson 1976, p. 164).

It may be possible for a control group within the experiment-proper to serve the function performed a manipulation check. For example, in the present study a control group was utilized of variable levels presented to measure acceptance as background information for all subjects.

Comparisons were made between the levels _perceived_ by the control groups and those reported by the experimental groups, rather than assuming that the subjects had perceived the levels as presented.

Choice of appropriate manipulation check measures. The sort of measure which is appropriate for a manipulation check depends on the type of variable which is involved and the kind of statement the researcher would like to make with regard to the independent variable when the experiment has been completed, that is, it depends on the _conceptual definition_ on which the _operational definition_ is based.

For example, if the independent variable were an event-type variable such as the number of exposures of an advertisement, then whether or not a manipulation check were even required would depend on how exposure was defined. If it was defined as the number of _presentations_ (stimulus-oriented), then unless there were a procedural error, the experimenter may be certain that the independent variable was, in fact, manipulated. However, if the conceptual definition were response-oriented, then a measure would be required. If exposure were defined as _paying attention_ to each presentation, then some measure of attention would be appropriate. If the independent variable were defined as the _perceived_ attractiveness or _perceived_ credibility of a spokesperson (both response-oriented variables) then the perceived attractiveness or perceived credibility should be measured.

The same principles hold for verbal manipulations. If the occurrence of the independent variable is not under the direct control of the experimenter (i.e., if it depends on a response by the subjects) then it must be measured before the experimenter can know that it has occurred.

Verbal manipulations may fail in two ways. First, the crucial part of the manipulation may be missed by the subjects, i.e., they may not attend to it. Second, even if subjects attend to the critical part of the verbal manipulation, they may not interpret it as the experimenter had intended (Fromkin and Streufert 1976, p. 447; see also Festinger 1953, p. 157). Both cases will result in a failure to manipulate the independent variable, but the remedy for the two problems is different. This first problem may be solved by making the manipulation more obvious--by calling attention to it is some way. The second problem requires changing the content of the verbal material.

In choosing what to measure in the manipulation check, the researcher may pick either the independent variable or variables known to be associated with it. Associated variables may be surrogate indicators of variation in the predictor variable or they may indicate surplus meaning. For example, it may be difficult to manipulate the perceived attractiveness of a communicator without also manipulating perceived credibility. If the researcher intended to manipulate only perceived attractiveness, then he or she should not only measure perceived attractiveness but also perceived credibility in order to assess the extent of extraneous variation.

In some cases the intended independent variable is a multi-dimensional concept. In such cases all of the intended dimensions

should be assessed to determine if the manipulation is deficient. For example, in the present study credibility was considered to be jointly defined by the concepts expertise and trustworthiness. Both of these factors were measured within the experiment, after the assessment of the independent variables (as were the associated concepts of believability and whether the communicator's position was justified). However, only the perceived expertise factor was significantly different between the high and low credibility conditions. Thus, the manipulation was deficient in effecting the twin dimensions of credibility and, in addition, the associated variables did not covary with the experimental conditions in the predicted manner.

In some cases the independent variable cannot be measured directly by asking subjects for a self-report. "...The major difficulty with the use of introspective reports; too often, subjects are unable or unwilling to explain just what the effects of some treatment have been.... People are not always in touch with their feelings, nor are they always able to articulate subtle psychological states" (Carlsmith, Ellsworth and Aronson 1976, p. 164).

Because of their relative transparancy, direct assessments of the independent variable may also be affected by the implicit social interaction between the subjects and the experimenter just as is sometimes observed for measures of the dependent variable. Therefore, an indirect measure of the hypothetical construct may be preferred. Possibilities are verbal measures of associated variables, measures of behavior, and psycho-physiological measures (see Carlsmith, Ellsworth and Aronson 1976, Chapter 2).

A question that must be confronted regardless of the type of measure used is whether the measure as operationalized has construct validity. The researcher should establish that the operation used in the manipulation check measures the intended variable. This is established by examining the relationship between measures of the intended variable by the tentative operation with measures of the same and other variables by other independent operations. The researcher's confidence in the tentative operation as a measure of the intended variable (i.e., in its construct validity) increases if the relationship between measures taken using the tentative operation and those taken using other operations is as predicted.

The predicted relationships are derived from the theoretical network of variables within which the target construct is embedded (i.e., nomological net (Cronbach and Meehl 1955)). The nomological net specifies which variables the target construct should be related to and the nature of their relationship.

At a minimum, measures of the construct generated using the tentative operation should be strongly, positively related to measures of the same construct generated using other, independent operations. The agreement in measures generated using different operations is called convergent validity (Campbell and Fiske 1959, p. 274) or trait validity (Campbell 1960, p. 547).

Since measures of the same construct generated using similar operations may share method variance as a confound, convergent validity is best established through "...attempts to measure the same trait through maximally different methods" (Campbell and Fiske 1959, p. 277; also see Cohen 1979, p. 305, for a discussion of this issue).

Campbell (1960, p. 547) distinguishes between trait validity and the validity established through the examination of relationships between measures of the target construct generated by the tentative operation and measures of other constructs. Campbell refers to this latter type of validity as <u>nomological validity</u> (p. 547). The predicted relationships examined in attempting to establish nomological validity may be positive, negative, or zero. Observed relationships between measures of the target construct and measures of other variables are compared with the predicted relationships to determine if they are as predicted by the theoretical framework depicted in the nomological net. Campbell and Fiske (1959) call this type of validation <u>discriminant validation</u> to emphasize that the tentative operation should <u>discriminate</u>, that is, it should not correlate highly with operations "...purporting to measure different things" (p. 277). However, it should be noted that the nomological net <u>may suggest high correlations</u> between some constructs. The reasearcher's confidence in the construct validity of his or her operationalization increases when the observed relationships are as suggested by the nomological net. However, as in so many other cases, negative evidence may be definitive whereas positive evidence may not. The reader is referred to Campbell and Fiske (1959) for examples of how correlations examined in the construct validation process may be arrayed in a <u>multitrait-multimethod matrix</u>.

<u>Interpretation of the results of a manipulation check: among- and within-condition variation</u>. In evaluating the differences among treatment means, the researcher should not forget the goal of the manipulation. The manipulation should achieve widely disparate levels on the intended dimensions and be relatively free of surplus meaning.

Therefore, statistically significant differences on the intended dimensions is a necessary, but not sufficient condition for concluding that the independent variable was successfully manipulated. The intended levels must also have been achieved. In other words, the differences between conditions on the manipulation check must be practically different as well as statistically different. This is a subjective judgement but it can be guided by examining the possible response categories on the manipulation check measure and applying a "reasonable person" criterion: Would a reasonable person, who had nothing to lose or gain through the publication of the data, agree that the intended levels of the independent variable had been achieved?

As an example, consider the operationalization of expectancy by Gorn and Goldberg (1977). Their subjects ("eight-to-ten year old boys," p. 86) were told either that they would have a one-in-fifteen, an eight-in-fifteen, or a fourteen-in-fifteen chance of winning a prize. These operations were to represent "low", "moderate" and "high" expectancy conditions. A manipulation check was performed prior to the presentation of the second independent variable (exposure to zero, one, or three commercials) and, consequently, before the assessment of the criterion variables.

Possible response categories were scored from one to five, with one being "lose for sure" and five being "win for sure" (p. 87). Means for the three expectancy conditions were: low = 3.06, moderate = 3.55, and high = 3.62 (on a 1 to 5 scale). They reported a significant main effect ($p < 0.05$) for the expectancy manipulation on this measure. While they acknowledge that "...the range of mean scores (3.1 to 3.6) was constricted toward the middle of the five point scale," they

attribute this as "...probably an artifact of the anchor points used for the scale" (footnote, p. 87). They proceeded with the analysis and reported "...no differences in behavior or attitude across expectancy conditions" (p. 88). In reading this one could conclude that expectancy failed to influence attitude or behavior. However, although their manipulation check revealed statistically significant differences between conditions, it appears that they succeeded only in achieving three levels of moderate expectancy. Rather than questioning their scale they should have questioned their manipulation. The analysis should not proceed if it is apparent that the independent variable was not manipulated, as the results will not be meaningful.

Another use of the manipulation check is the assessment of the within condition error variance. Although there is no clear standard for determining if the error variance is "large" (except through comparing it with the differences among condition means), the distributions of individual scores for the conditions should be examined for overlap. If there were conditions _labeled_ low, moderate, and high, it is possible that a person, say, in the "moderate" condition could receive an effective manipulation that was higher than that received by an individual in the "high" condition. Such overlap should be avoided if possible.

The implications of manipulation check results. If the analysis of manipulation check results indicates that the conditions are statistically and practically different, that the within condition variance is not "significantly" large, and likely artifacts do not vary significantly between conditions, then the analysis may proceed and the relieved researcher may state his or her conclusions with greater

confidence and a clear conscience. An additional step would be to effect other manipulations of the same independent variable(s) within or across experiments to assess the generality of the observed effects.

If the analysis indicates that the manipulation resulted in surplus meaning, was deficient, or resulted in too much variability within conditions, then the researcher should revise his operations to achieve the desired effects. However, even if the manipulation check was run in the experiment, the effort is not necessarily wasted. Two possibilities exist. One is called secondary analysis, "...that is analysis in a new conceptual frame--which may give it a very different and often much greater significance" (Kaplan 1964, p. 160--Here Kaplan is referring to theoretical, rather than statistical significance, as the differences between means for the criterion variable would remain the same). This secondary analysis could take place if the manipulation check data implied that another unintended, variable had been manipulated rather than the one that was planned. In this case, subjects would have still been assigned randomly to conditions and there would still have been an experiment--it just wouldn't have been the one that was planned.

Another possibility is internal analysis. Internal analysis examines the relationship between the manipulation check measures and the criterion variable(s). It amounts to examining what the results might have been had the manipulation worked and can provide useful information for future experiments. However, "the fact that the experimenter once assigned subjects to treatments at random is irrelevant; the current analysis is not based on that assignment, but rather on a measured subject variable, and thus it is a correlational study" (Carlsmith, Ellsworth and Aronson 1976, p. 144). Before any

causal conclusion may be reached, the operations to effect the independent variables must be refined and another experiment run.

Persistent Sources of Experimental Artifacts

The rest of this chapter will be devoted to a discussion of several sources of experimental artifacts that should receive special attention in the planning and conduct of an experiment. These are the experimenter, the subjects, the instructions, and any questionnaires which may be administered.

The Experimenter as a Source of Experimental Artifacts:

Research has dispelled what Friedman called "...the democratic notion that all experimenters are created equal; that they have been endowed by their graduate training with certain interchangeable properties; that among these properties are the anonymity and impersonality which allow them to elicit from the same subject identical data which they then identically observe and record" (1967, pp. 3-4). Kintz et al. (1965) cite evidence that experimental outcomes are influenced by the experimenter's personality, perceived personality, experience, sex, expectancy, skill at the experimental task (modeling), and awareness of the performance of the initial subjects (early returns). They also cite evidence that the experimenter's personality interacts with that of the subjects' in its influence on the outcomes. (See also Rosenthal 1977, p. 257, and Carlsmith, Ellsworth and Aronson 1976, p. 168). These factors may be a source of error variance or they may be a source of systematic effects (through confounding and/or interactions).

The process by which personological variables influence experimental outcomes is not clear. "The sex, age and race of the investigator have all been found to affect the results of research. What we

do not know and what we need to learn is whether subjects respond differently simply to the presence of experimenters varying in these biosocial attributes, or whether experimenters varying in these attributes behave differently toward their subjects and, therefore, obtain different responses because they have, in effect, altered the experimental situation for their subjects" (Rosenthal 1977, p. 256).

Whatever the role of personological variables in affecting the criterion measures, it is clear that experimenter behavior varies both across experimenters and within experimenters across subjects. Friedman and his associates studied the behavior of twenty-nine experimenters whose experimental performance had been recorded on film. "Psychological experimenters are supposed to be inflexible, mechanical, 'programmed', standardized in their behavior. These experimenters improvised and ad-libbed and were nonconforming, different, variable in their behavior" (Friedman 1967, p. 106).

Friedman points out that many aspects of the experimenter's behavior have no established standards. "...How should the subject be greeted? How far apart should the experimenter and the subject be seated? Should the experimenter ever look or smile at the subject? How long should the experimenter take in reading the instructions? In what tone of voice should different parts of the instructions be read? At the conclusion of the experiment, how should the subject be dismissed?" (Friedman 1967, p. 73).

His observations show how variable even the beginning stages of an experiment can be. He observed the following statements for experimenters who were supposed to read just the word name: "Name? Your name? Your name, please. What is your name? Name, please. Can I have

your name? May I have your name, please? I want to get some information about you. First, your name. What's your name? Can I have your name, please? First name? Last name; first name. Let's have your name, please. Your name, miss? Now, I'd better get your name" (pp. 81-82). Experimenters were also observed to vary the instructions, the presentation of materials, and the labeling of scale end-points (1967, p. 97).

Friedman also observed differential experimenter behavior that appeared to result from an interaction of experimenter and subject variables. "Male experimenters are more likely to smile when greeting female than male subjects, and are more likely to use a female's name... or an affectionate equivalent... Here, then, we have a new sex difference to add to the literature: the sex of the subject affects the behavior of the experimenter. The question is: once we know this can we quote from the research literature on sex differences quite as confidently as we did before" (Friedman 1967, pp. 104-105).

Controlling Experimenter Artifacts:

There are several ways of controlling experimenter artifacts, and all follow the principles of control outlined earlier in the chapter. The control of personological variables may be controlled through random assignment of subjects to experiments (Carlsmith, Ellsworth and Aronson 1976, p. 16); through balancing (having each experimenter run an equal number of subjects in each condition (Keppel 1973, p. 523); through introducing experimenters as an independent variable in the design; or through the use of one experimenter for all conditions.

Experimenter expectancy effects are probably best dealt with by keeping experimenters "blind" to the subjects' conditions, or at least,

in a factorial design, by keeping the experimenter blind to the level of one of the variables, i.e., "half-blind" (Carlsmith, Ellsworth, and Aronson 1976, p. 276). Also running all conditions simultaneously with the experimenter naive as to the individual subject assignments will control expectancy effects.

Variability of behavior between experimenters could be solved by using one experimenter and running all subjects simultaneously. However, this may limit the generality of the observed effects. "Standardization" of experimenter behavior may be achieved through training and observation. Experimenters should run experimental conditions in pilot studies so that they can become more familiar with the procedures and develop more consistant behavior. Experimenters should also be appraised of the effects of variability in their behavior in terms of possible bias and the inflation of the error term.

Experimenters may also be observed during the conduct of the experiment. Experimental sessions may be tape recorded, video-taped, or observed through one-way mirrors. Stooges may also be used as "subjects" as a check on experimenter behavior (Friedman 1967, p. 108).

Generality of the findings is increased through the use of heterogeneous experimenters. By including experimenters as an independent variable, then main effects of the experimenters will be removed as a source of bias and any remaining biasing effects may be detected by the presence of an interaction between the primary independent variable and the control variable, <u>experimenter</u> (Keppel 1973; Peterson 1977 is an example of controlling for experimenter effects in this fashion).

Subject Variables as a Source of Experimental Artifacts:

Personological variables:

Many personological variables may influence the criterion variable. Intelligence, sex, education, past experiences related to the criterion, group memberships, age, etc. may all affect the criterion variable. Therefore it is essential that these variables be controlled. The simplest way of controlling these variables is through randomization, i.e., through random assignment to conditions. This should also help control for the interactive effects of subject variables eliciting differential behavior from the experimenters. However, many other subject effects are not so readily dispensed with. These effects are not inherent subject variables, but are unique to the experimental setting. They arise from the subjects' exposure to and response to aspects of the experimental situation.

Voluntarism. One subject variable which has received considerable attention is volunteering for an experiment (Rosenthal and Rosnow 1975). "...the overall evidence is suggestive of such differences as: Volunteers are higher in need for social approval. Volunteers, especially males, are more intelligent. Volunteers tend to have more unconventional personalities. Volunteers more often tend to be first-borns. Volunteers for certain experiments are less well adjusted. Volunteers tend to have higher need for achievement" (Jung 1971, p. 28).

The primary problem with voluntarism is not with internal validity (provided volunteers and coerced subjects are not confounded with experimental conditions), but with the generality of the experimental effects which are observed for volunteer samples. Would the same effects be observed for subjects who were not volunteers? The question is concerned

with whether the independent variable interacts with voluntarism. However, Kruglanski points out that it may be misleading to discuss voluntarism as a singular variable. "...Rather than signifying a convergence of several operational definitions upon a unitary psychological construct, the notion of volunteer status seems to denote the convergence of numberous constructs (notably the various motives for volunteering) upon a single act or operation: the showing- or signing-up for a psychological experiment. Such convergence appears to be of little scientific interest" (1975a, pp. 112-113).

Kruglanski also points out the relevance of Bakan's (1955) notion of general- versus aggregate-type research to the problem of voluntarism "...in general-type research the proposition at stake is expected to apply to each and every member of the designated class. Hence, in such research any member or a group of members is sufficiently representative" (1975b, p. 380). In fact, for theoretical research, there is rarely any clearly defined target population to which the results are to be generalized. The problem or representativeness would be relevant in applied, or problem oriented, research in which estimates of parameter values for the population are desired (i.e., in aggregate-type research).

However, even in cases for which representative samples were desired, randomly selected individuals would not be all equally willing to participate. It would be difficult to establish to what extent subjects had in fact "volunteered" and to what extent they felt coerced into participating. Most decisions to participate probably reflect at least some form of normative pressure. As such, voluntarism is probably best viewed as a continuous, rather than as a categorical variable.

Many university subject pools (nee research participant pools) offer the members some abhorent alternative to experimental participation (volunteering). In the present study, subjects were solicited from intact groups (social organizations). Their organizations received monetary payment for each member's participation. Because of likely peer-pressure to participate, it is difficult to discern just what "volunteering" means in such a situation. In a sense, anyone who agrees to participate in an experiment is a volunteer.

Rosenthal and Rosnow (1975, pp. 198-199) point out several methods of controlling for voluntarism as a threat to the generality of experimental findings. They consist of making the appeal for volunteers in such a way that volunteering is more attractive to the potential subject. The experiment should be presented as interesting, non-threatening, important, and non-stressful. Volunteering should be presented as responsible and as normative for the target population. The appeal should be made by someone of high status (preferably a woman) and also someone known to the target population. Potential subjects should be asked for a public commitment of their willingness to participate if it is felt that they perceive that volunteering is, indeed, normative for their population. Potential subjects may even be given token gifts for their time spent considering participation.

If Rosenthal's and Rosnow's suggestions are successful, then the class, volunteers, will be broadened to include a more hetergeneous sample, thereby increasing the generality of the findings from studies in which such volunteers participate.

However, the only way to truly control for voluntarism is to make it unnecessary--for the subjects to be unaware that they are

participating in an experiment. However, there are serious ethical implications in such an approach.

Naivete versus sophistication:

The ideal situation in an experiment would be for subjects to respond to the experimental treatment in a natural, guileless fashion. However, except in cases where the measurement and manipulation are unobtrusive, the subjects are aware that they are participating in an experiment. This awareness can lead to responses which have serious implications for the relational statements made by the researcher after the experiment. Subjects are not passive, trusting creatures who wait patiently for the experimenter to do what he or she will with them. "Simply stated, people feel and think, and people who are subjects in psychological experiments are undoubtedly feeling and thinking about the experiment" (Silverman 1977, p. 2).

Awareness of and thinking about being in an experiment can lead to suspicions about what is really going on (see Stricker, Messick, and Jackson 1969, p. 348). This suspiciousness is probably heightened in situations in which subjects are likely to be aware of the frequency with which deception is used. This awareness is spreading beyond the academic setting as students who have participated in deception studies graduate and mass media events such as the televising of the Milgram (1963) study spread the word.

Hypothesis guessing and demand characteristics:

The suspicion aroused in an experiment can lead to hypothesis-guessing. The effects of hypothesis guessing on the experimental outcome have been labeled demand effects and the cues that influence subjects' guesses are called demand characteristics. Demand

Characteristics are cues which "...include the rumors or campus scuttle-butt about the research, the information conveyed during the original solicitation, the person of the experimenter, and the setting of the laboratory, as well as all explicit and implicit communications during the experiment proper... [They include] the experimental procedure itself" (Orne 1961, p. 16-17).

Closely related to the concept of demand characteristics is the concept of <u>reactive arrangements</u> (Campbell 1957). The concept of reactive effects is essentially the same as an interactive effect between aspects of the experiment and the treatment. These interactions create"...reactions to <u>X</u> which would not occur had <u>X</u> been encountered without this 'I'm a guinea pig' attitude... The problem of reactive arrangements is distributed over all features of the experiment which can draw the attention of the respondent to the fact of experimentation and its purposes" (Campbell 1957, p. 279). Campbell's <u>reactive arrangements</u> appear to have initially been more general than demand characteristics, dealing not simply with hypothesis guessing, but with <u>any</u> aspect of the experiment may interact with the treatment to make the subject respond differently to the treatment because of being in the experiment. However, it appears that currently there is little difference in the concepts. "...The concept of demand characteristics has been broadened to designate all cues in the experiment that may influence the subject" (Adair 1973, p. 24).

<u>Assessing the impact of demand characteristics</u>. Orne has outlined several methods for assessing the impart demand characteristics. The first overall approach consists of asking subjects what they think the hypotheses were. This may proceed after the conduct of the experiment

("postexperimental inquiry") or may follow a description of the experimental treatment ("pre-experimental" inquiry). A modification of the post-experimental inquiry is to run "...different sets of subjects to different points in the experiment, stopping at these points as if the experiment were over (for these subjects it, in fact, it is over) and [carry] ...out inquiries" (Orne 1969, p. 155).

The post-experimental inquiry suffers from several shortcomings. One is what Orne called "...a pact of ignornance resulting from the intertwining motives of both experimenter and subject, neither wishing to create a situation where the particular subjects's performance needs to be excluded from the study" (Orne 1961, p. 18).

Another shortcoming of the post-experimental inquiry is "...that the subject's perception of the experimenter's hypotheses is based on his own experimental behavior, and therefore a correlation between these two variables may have little to do with the determinants of behavior" (Orne 1961, p. 18).

An additional weakness of the post-experimental inquiry "...is that it places the subject in a cognitive set possibly quite different from that present during the experiment proper" (Kruglanski 1975a, p. 129).

Finally, "...the inquiry procedure itself is subject to demand characteristics" (Orne 1961, p. 18). "While these [posttests] may sometimes demonstrate a method to be reactive, they may fail to detect many instances in which reactivity is a serious contaminant. Subjects who consciously dissemble during an experiment may do so afterward for the same reasons. And those who are unaware of the effect on them at the time of the research may hardly be counted on for valid reports afterwards" (Webb *et al.* 1966, p. 16).

In the pre-experimental inquiry (or non-experiment), the experimental treatments are described to the subjects, but the actual treatment is not presented. Subjects are asked to participate in a type of role-playing. "Thus one would say, 'If I had asked you to do all these things, what do you think that experiment would be about, what do you think I would be trying to prove, what would my hypotheses be?' etc...[in addition, subjects may be given]...post experimental tests. The subject is requested to behave on these tests as though he had been exposed to the experimental treatment that was described to him" (Orne 1961, pp. 18-19).

Orne suggests that the convergence of the results of post-experimental tests of role playing subjects and actual experimental subjects is evidence that the experimental results could be accounted for by demand effects. However, convergence doesn't establish either the demand characteristics or the independent variable(s) as the singular cause of experimental results, it just allows <u>both</u> to remain as possible causes. Only when the pre-experimental inquiry yields <u>different</u> results from the experimental results may demand characteristics be rejected as an alternative explanation of the experimental findings (Orne 1969, pp. 157-158).

Kruglanski and Eilam point out other alternative explanations for the convergence of pre-experimenter inquiry and experimental results besides the possibility that experimental results are caused by demand characteristics. First, the process operating in the response to the description of the treatment may be the same as that operating in the response to the treatment. Second, subjects may be able to successfully predict response to the treatment because of an awareness of the

relationships, perhaps because of experience or empathy. Third, "...adoption of an intense problem solving set by the role-playing subjects [may enhance]... their likelihood of zeroing in on cues disclosing the experimental hypothesis" (Kruglanski and Eilam 1974, p. 15). "The ...multiplicity of possible interpretations suggests that the role playing techniques for detecting 'demand characteristics' raise more problems than they intend to solve. Personally, we think a thorough and skillfully conducted post-experimental interview seems still the most efficient method for identifying the artifacts inherent in given experimental procedures" (Kruglanski and Eilam 1974, p. 15).

Control of demand characteristics. Several approaches have been suggested to control for demand characteristics. The most extreme is "disguised experiment". In this approach the subjects, and possibly the experimenter(s), do not know that they are involved in an experiment (Campbell 1969, p. 368-369). However, there are procedural and ethical problems associated with this approach.

The second, and most common approach is the use of deception. The idea is to give the subjects a common, yet false, hypothesis to which they may respond, rather than allowing everyone to form his or her own hypothesis. This should reduce bias and the error variance in the experiment, but may limit the generality of the findings. However, the decline of the naive subject makes it unlikely that all subjects will be taken in by the false hypothesis (Kelman 1967). As a result, the desired uniformity of supposed bypotheses across subjects will probably not be achieved. "...The use of deception, while it is designed to give the experimenter control over the subject's perceptions and motivations, may actually produce an unspecifiable mixture of intended and unintended

stimuli that make it difficult to know just what the subject is responding to" (Kelman 1967, p. 6).

A third approach utilizes a false hypothesis but also strives for "experimental realism". An experiment has experimental realism "...if the situation is realistic to the subjects--if they believe it, if they are forced to attend to it and take it seriously--in short, if it has impact on them" (Carlsmith, Ellsworth and Aronson 1976, p. 81).

"If subjects are deeply involved in the experimental situation, or interested in it for its own sake, they will be less likely to turn their attention to other things, such as conjectures about the experimenter's intentions. The subjects should care more about the apparent nature of their immediate experience than about their ability to assess its true nature..." (Carlsmith, Ellsworth and Aronson 1976, p. 121).

Experiment-relevant subject motivation:

Given that subjects may speculate about possible hypotheses in an experiment, the next question is: What are they going to do about the guesses that they make? "The subject artifacts that have captured the particular attention of investigators were all related to the subjects' goals or motivations with respect to the laboratory situation" (Kruglanski 1975a, p. 111; see also Rosnow and Davis 1977, p 303). The motivations that are aroused arise from the social interaction between the subject and the experimenter and between the subject and other subjects participating in the experiment.

Motivations arising from the social interaction between the subject and the experimenter. In the social interaction between the experimenter and the subject there are several roles that the subject may take. The idealized role of the subject pictures him or her as a naive

creature who will follow directions carefully, will understand the manipulations as intended by the experimenter, will respond guilelessly and naturally to the treatments with a lack of concern for the impact of that response on the image of him or her it may portray to others.

However, social interactions provide opportunities for mutual exploitation between the subject and experimenter, with each jockeying for outcomes favorable to himself. The particular role that the subject chooses will influence the responses that he will make to any perceived demand characteristics (Weber and Cook 1972, p. 276).

"In sociological role theory the basic assumption is that human performance is affected by the social prescriptions and behavior of others. Social behavior is viewed as a response to certain implicit demands associated with specific propriety norms and of various mediating factors that may make people more or less compliant with role expectations" (Rosnow and Davis 1977, p. 303).

Several possible roles have been discussed in the literature. The one which most closely approximates the idealized subject is the "faithful subject" (Fillenbaum 1966). "The faithful subject is someone who believes that a high degree of docility is required in research settings and who further believes that his major concern should be to scrupulously follow experimental instructions and to avoid acting on the basis of any suspicions he might have about the true purpose of a study" (Weber and Cook 1972, p. 275). This role is not a source of bias. This is the same as the role of "honest subject" suggested by Orne (1961, p. 17). Orne expected that this role would be most likely when the hypothesis was <u>obvious</u>. But, then, Orne was prone to attributing benevolent motives to subjects as in his "good subject" designation

(1961). "The essence of the good subject is that he attempts to give responses that in his opinion, will validate an experimental hypothesis" (Weber and Cook 1972, p. 275). This was the role that most concerned Orne (1961).

Another role which concerns researchers is that of a subject who attempts to manage the impressions he makes on the experimenter. This involves the subject's determining his response through a sort of "negotiation.... The subject can both inhibit a behavior and can stimulate it, depending on how he believes a particular kind of behavior will be judged by the experimenter" (Riecken 1962, p. 101-102). Carlsmith, Ellsworth and Aronson consider this impression managing role to be the primary role adopted by experimental subjects. They suggest that the goal of hypothesis guessing on the subject's part "...is a desire to find out what behaviors are expected of the good, healthy, intelligent, and normal person so as to adjust his or her own behavior accordingly..." (1976, p. 282). Closely allied to this role is the role of the "apprehensive subject" (Rosenberg 1969). Apprehensive subjects "...are apprehensive about how their performance will be used to evaluate their abilities or their socioemotional adjustment" (Weber and Cook 1972, p. 275).

Another role which subjects may adopt is that of the "negativistic subject" (Cook et al. 1970). The negativistic subject "...is assumed to want to infirm it [the hypothesis] by corroborating some hypothesis other than the experimenter's or by giving responses that are of no use to the experimenter" (Weber and Cook 1972, p. 275). There are several reasons why the subject may adopt this role. Rosnow and Davis suggest the applicability of Brehm's (1966) reactance theory to subject's

perceptions of their situation with respect to the experimental hypothesis. "...We postulate that when cues are so patently obtrusive that they tend to restrict the subject's freedom of movement, this will have a strong dampening effect on positive motivation" (Rosnow and Davis 1977, p. 305).

Another negative motivational influence arises not from the treatment, but from having to participate in the experiment. Many university subject pools use thinly veiled coercion to ensure the participation of an adequate number of "volunteers". Usually, some odious alternative to participation is offered the subjects. In an effort to re-establish equity, i.e., "get even", subjects may elect to sabotage the experiment. Thus, coercion may reduce the negative effect on generality resulting from voluntarism, yet at the same time result in negative motivational effects which reduce the meaningfulness of the results.

Task-oriented motivational influences. The role which is adopted by the subjects is probably influenced in part by the experimental task which they are asked to perform. As mentioned earlier, subjects whose freedom-of-response options is limited by the demand characteristics implicit in a task may lead to negativistic responses. Fromkin and Streufert suggest that evaluation apprehension may be reduced if subjects have the perception that "...the purpose of the experiment is more of a mathematical or technical nature than an investigation of subjects' personalities..." (1976, p. 438).

Tasks also vary in what could be viewed as good behavior (Riecken 1962, p. 100-102). Tasks that may be perceived as measuring task ability (e.g., ability, skill, or capacity) have no upper limit in what

may be viewed as "good" behavior. If the effect of some non-motivational variable (e.g., some facilitating variable) on performance is being assessed by the experimenter then demand effects do not threaten the ability of the researcher to draw conclusions from the results (Orne 1969, p. 156). In such cases motivation to do well by all subjects should increase the internal validity of the experiment--it should increase uniformity of motivational effects both across and within conditions, thereby reducing both bias and error variance.

For other types of tasks, there may be a midscale value which would be considered good (Riecken 1962, p. 101) in which case motivation to do well would tend to reduce differences in effects between conditions. In other cases, the ideal level may be unknown to the subjects, in which case motivation to do well on the task would probably just increase error variance.

Another possibility is that the performance level which is perceived to be "good" may be confounded with treatment levels. A minipulation check should be performed to assess this possibility.

Other aspects of the task that may be expected to influence subject motivation are the minimum ability level required to perform the task and how inherently interesting the task is. It is easy for a researcher who has been studying a topic for months to design a task that makes perfect sense to him or her, yet is incomprehensible or frustratingly complex for the subjects (see Richey 1976). If the task is too difficult, or if it is intrinsically boring, then subject motivation will decrease and error variance and/or bias may increase.

Motivations arising from the social interaction between subjects. Not only is there social interaction between the subject and the experimenter, there is also interaction between subjects. This interaction can occur within and between experimental sessions.

When subjects are run in groups they can overtly (or covertly) verbally and nonverbally share information regarding their interpretations of the situation and what appears to them to be appropriate behavior in the situation.

Therefore, if one subject learns the hypothesis, all may learn it. All subjects may also share an erroneous hypothesis. Impression management may take place between subjects when they know that other subjects will be aware of their responses.

Modeling behavior may also occur. If subjects are unsure of how to complete the task they may observe others and mimic their behavior. This is probably expecially true of the apprehensive subject who wants to be sure to "do the right things".

Inter-subject influences may also occur between experimental sessions as veteran subjects share information with "naive" subjects whose turns have not yet come.

A summary the of subject variables on experimental results:

To review, subject variables may affect experimental results in three ways: (1) subject variables confounded with treatments may artificially inflate or conceal differences between treatment means either through their main effects or through interaction with the treatment variable; (2) subject variables which are constant across conditions may limit the generality of research findings; and (3)

subject variables whose impact on the criterion variable fluctuates randomly within conditions may inflate the error variances such that true differences in treatment effects will not be detected.

Thus, hypothesis guessing could only bias experimental results when the hypotheses guessed or the tendency to generate hypotheses vary systematically with the experimental treatments (Weber and Cook 1972, p. 280). The same is true for subject motivations to respond to their guesses.

"The effects of suspicion on behavior in an experiment are apt to be extremely varied, depending upon the subjects' perceptions and motivations connected with the study.... With the many possible differences in the perceptions and motivations of suspicious subjects, the only certainty is that their disblief increases extraneous variation in a study" (Stricker, Messick and Jackson 1969, p. 347).

When the experimental design is complex, with several independent variables with several levels each, it is unlikely that even subjects who successfully guessed the hypotheses could successfully fake behavior consistent with (perhaps non-obvious) main effects, much less interactive effects. "Complying with a differential hypothesis is a complex process which requires that the subject knows, first, what the other experimental conditions are ; second, how persons in each of these conditions are supposed to behave; and third, how he can calibrate his hypothesis--confirming behavior so that he registers in responses the hypothesis that he learns in words" (Weber and Cook 1972, p. 277).

In an extensive literature review Kruglanski (1975a) concludes that "...the findings ...provide little support for the suppositions (a) that the subjects generally form systematic hypotheses contingent on the

experimental treatments, (b) that any given subject's motivation (e.g., cooperativeness, negativism, faithfulness, or apprehensiveness) is typically aroused within the subject's psychological research, and (c) that the volunteers for experiments are characterized by any consistent set of psychological qualities.... The conception of the 'laboratory situation' seems to lack internal consistency as a seat of the various subject artifacts" (Kruglanski 1975a, p. 141). However, as Kruglanski points out, the fact that these artifacts are not routinely produced by the laboratory situation does not mean that they would not affect the experimental results on those occasions on which they might occur. However, "...the only conditions that seemed to produce pervasive effects were those of (1) providing a hypothesis to the subjects, and (2) cueing them as to what the typical or socially desirable response would be" (Kruglanski 1975a, p. 141).

Control of Subject Artifacts:

Artifacts arising from the subject-experimenter interaction:

The control of artifacts that may arise from the social interaction between the subject and the experimenter may be achieved by reducing insightful guessing of hypotheses and associated behaviors and by the control of subject motivation vis a vis the experimental situation (Rosnow and Davis 1977, p. 308).

Concealing the true hypothesis. The control of hypothesis/behavior guessing may be effected by camouflaging the experimental hypothesis and by the avoidance of cues which would suggest appropriate behaviors to the subjects (see Weber and Cook 1972, p. 291 and Kruglanski 1975a, p. 141).

The true experimental hypothesis should be made difficult to discover and a false hypothesis may be presented to the subjects. This false hypothesis should generate behaviors uncorrelated with those anticipated for the experimental hypothesis.

Controlling subject motivation. Several approaches have been suggested for controlling subject motivations. One is to purposefully cue the subjects to the faithful subject role (Weber and Cook 1972, p. 293). Since the faithful subject role is not a source of bias, cueing all subjects to this role should increase the internal validity of the experiment. Weber and Cook suggest that such cueing might be accomplished through "a serious task orientation, small groups of subjects who think their individual but anonymous responses are needed, [and] preliminary reference to the importance of a project..." (1972, p. 293).

Weber and Cook (1972) also suggest specific steps which may be taken to reduce apprehensiveness on the part of the subjects: "Experimenters should not be high-status persons. They should not control rewards or punishments. The more tangible sources of reward power (credit, payment, etc.) should be signed away before the experiment begins. Experimenters should not, explicitly or implicitly, claim skills in evaluating abilities or adjustment. Furthermore, the experiment should be designed so that all responses are at least anonymous. Experiments should also aim, wherever possible, to have subjects perform experimental tasks in settings where they cannot look to the experimenter for performance feedback" (1972, pp. 292-293).

Fromkin and Streufert (1976) add these suggestions for reducing apprehensiveness: subjects should perceive that "...(1) the purpose of the experiment is more of a mathematical or technical nature than an

investigation of the subjects' personalities; and (2) the experimenter is <u>not</u> interested in individual responses but in normative or nomothetic aspects of responses from groups of individuals" (p. 438).

Another approach to reduce spurious motivational effects is to maximize the experimental realism and the impact of the the setting and operations on the subjects (see the previous discussion p. 53, Carlsmith, Ellsworth and Aronson 1976, p. 121).

<u>Artifacts arising from the subject-subject interaction within experimental sessions</u>:

The control of subject artifacts which may arise from inter-subject interaction within an experimental session essentially involves restricting the flow of hypothesis-behavior-relevant information between subjects. This may be accomplished by controlling the opportunity and the motivation for communication.

<u>Restricting communication opportunities</u>. Within an experimental session information flow may be controlled in several ways. One way is through isolation. Subjects may be isolated so that the interchange of auditory and visual cues between subjects is restricted. In the present study subjects were isolated at stations constructed with the use of cloth-covered partitions. However, oral communication was possible and, in fact, some subjects were observed conversing. Some even moved their chairs so they could see each other while waiting for others in the group to complete intermediate tasks in the experiment. Another way of controlling for within-session, inter-subject information flow is to run subjects individually. However, this is very time consuming and allows for between-session information sharing.

Controlling motivation for communication. The desire for subjects to communicate with each other may be reduced by recruiting subjects for each session form heterogeneous sources (Carlsmith, Ellsworth and Aronson 1976, p. 306). The subjects within each session of the present study held common group membership.

Another way of minimizing information flow between subjects is to appeal to the "faithful subject". Subjects can be told of the importance of giving their own responses and the fact that the influences of others within the session on their behavior would reduce the meaningfulness of their participation.

Artifacts arising from subject-subject interaction between experimental sessions:

Controlling for inter-session contamination consists of limiting the opportunity and motivation for treatment-aware subjects to share information with naive subjects who have not yet participated, as well as limiting the information which aware subjects have.

Restricting communication opportunity. Inter-session opportunity may be eliminated by running all conditions simultaneously. This has the added advantage of holding time constant across conditions. However, if all conditions cannot be run in the same room, then experimenter will be confounded with the independent variable, as will any aspect of the room which varies across rooms.

Another way of limiting inter-session opportunity is to complete the whole study as quickly as possible (Carlsmith, Ellsworth and Aronson 1976, p. 306).

Both opportunity and motivation to share information may be reduced by recruiting subjects from heterogeneous sources (Carlsmith, Ellsworth

and Aronson 1976, p. 306). This would reduce the chance of inter-session encounters and the willingness to share information should such an encounter occur.

Controlling motivation for communication. Aronson (1966, p. 19) suggests appealing to the "faithful subject." He advises a thorough debriefing in which the harmful effects of inter-session contamination are stressed and the cooperation of the subject in limiting contamination is solicited. To reduce social pressure on the treatment- (and hypothesis-) aware subject to share information with future participants, they are provided with a limited description of the experiment which may be shared with others. This was the procedure followed in the present study.

Restricting available information. A last way of controlling intersession information flow is to limit the information given to experimental session participants. This requires delayed debriefing so that all subjects will be run before the experimental hypothesis is revealed. This may be accomplished when all subjects are accessible at a post-session time so that all may be debriefed at onece (Stricker, Messick, and Jackson 1969, p. 349). However, there are ethical problems associated with delayed debriefing. Maintaining any deception used in the experiment over a protracted time period violates the trust of the subjects and may in some cases be harmful to them.

Instructions as a Source of Experimental Artifacts:

The instructions given to experimental subjects are a crucial element in the experimental procedure. The instructions must "set the stage" (Carlsmith, Ellsworth and Aronson 1976) for the experiment; they must gain subject involvement; they must avoid providing the

experimental hypothesis to subjects; they must also avoid suggesting appropriate responses to subjects; and they should be be standardized to prevent contributing to the error variance.

Setting the stage:

"Setting the stage" is a term suggested by Carlsmith, Ellsworth and Aronson (1976) for the role of the instructions in providing the framework of the experimental drama within which subjects may conceptually operate during their participation in the experiment. "It is highly desirable to have some plausible and understandable purpose for the experiment which the investigator can communicate to the subjects and which they will accept. If this is not done, the subjects usually conjecture about it and make guesses as to the true purpose. If a plausible explanation is not given, this important aspect remains uncontrolled" (Festinger 1953, pp. 156-157).

The explanation of the purpose of the experiment and of the context within which the subject will operate is called a cover story. The presentation of the experimental treatment, the behavior of the experimenter and of other subjects, as well as the assessment of the dependent variable must be integrally related within the cover story and yet be presented in a fashion that will conceal the true hypothesis (see Fromkin and Streufert 1976, p. 442).

An essential part of the cover story is the false hypothesis presented to the subjects (i.e., deception). Any responses consistent with the false hypothesis should be orthogonal to the predicted responses under the true hypothesis. "Thus, it can be immaterial whether or not the subject completely accepts the false hypothesis (offered by the experimenter, so long as it prevents speculation about

the true experimental hypothesis" (Carlsmith, Ellsworth and Aronson 1976, p. 284).

Nevertheless, it is probably a good idea to assess in a pilot study whether subjects accept the false hypothesis. If the deception is ineffective, then either error variance will increase as subjects generate their idiosyncratic hypotheses or they may perceive the true hypothesis. "If the dissimulation is not perceived as authentic, the subjects' skepticism will be aroused. Hence, suspicion is a useful index of the effectiveness of deception" (Stricker, Messick and Jackson 1969, p. 344). Post-experimental inquiries (as discussed earlier, p. 51) are a useful way of assessing the success of deception and the level of suspicion.

In setting the stage, subjects may also be appraised of appropriate protocol within the experimental setting. Thus, from the start subjects should know whether communicative exchanges with other subjects or with the experimenter are appropriate. They should know what task-relevant behaviors are expected of them and what they may expect of the experimenter.

Gaining subject involvement:

Lack of subject involvement with the experiment can have several serious consequences. Uninvolved subjects will not be motivated to listen to instructions or to try to complete the tasks as intended; they may play games in trying to outsmart the experimenter or to sabotage the experiment. To the extent that these responses are uncorrelated with treatment conditions, they will be a source of error variance and may obscure any real treatment effects. If the responses are confounded with experimental treatments, then they may either obscure or falsely

indicate treatment effects. To gain subject involvement, the situation presented to the subjects should be "cognitively 'real'" (Festinger 1953, p. 153), that is, it should be high in what Carlsmith, Ellsworth and Aronson (1976, p. 81) call "experimental realism" (see the earlier discussion, p. 54).

"The situation encountered by subjects in an experiment must be so striking and so believable that its effects will transcend the influence of their knowledge that they are in an experiment" (Carlsmith, Ellsworth and Aronson 1976, p. 83).

Avoidance of hypothesis-relevant cues:

The instructions must make clear to the subjects exactly what behaviors are expected of them in performing experimental tasks. Instructions must do this in a manner that is so involving that subjects willingly participate, yet at the same time must avoid providing any cues which will reveal the hypothesis or hypothesis-relevant behaviors.

If the context within which the treatment is presented is too simplistic then the alternative hypotheses which subjects may guess may be reduced to the point that, guessing the true hypothesis may be easy. For example, just saying to subjects: "We are going to show you some commercials and then ask you some questions about them," allows them to restrict their search for possible hypotheses to aspects of the commercials (see Baker and Churchill 1977 for an example of a similar presentation). However, if the commercials are presented within a broader context in which they form a natural part of the overall script (for example imbedded in a television program for which a plausible reason for viewing has been established) then the true hypothesis should be more easily concealed.

With increasing complexity, impact and concealment of the experimental hypothesis should increase, but "if we gain impact by elaborating the treatment to make it striking and complex, the experimental effect may be caused by any one of the elements we have added rather than by the 'pure' independent variable we had in mind" (Carlsmith, Ellsworth and Aronson 1976, p. 61).

Researchers should also be careful to avoid cueing subjects to hypothesis-relevant behaviors by suggesting certain modes of response in the instructions. For example, in Chapter Three, the instructions used by Petty, Wells and Brock (1977) were discussed. They were examining the effects of distraction on the level of negative subvocal responses reported on a post-message verbal protocol. However, in their instructions to subjects, they suggested that subjects might want to write down negative or positive statements or a mixture of both. By suggesting the evaluative mode of response the researchers may observe effects that could not be produced without the possibly reactive instructions, thus limiting the generality of the findings.

Instruction standardization:

If instructions are not standardized across subjects, then as a source of error variance or through correlation with the treatments, variations in presentation could lead to falsely rejecting the null hypothesis or could produce spurious differences in treatment means. Friedman (1967) revealed many variations in instruction presentation across experimenters and for individual experimenters across subjects. As discussed earlier under Controlling experimenter artifacts (p. 45) standardization of instruction presentation may be improved through training and observation, and by impressing on experimenters the significance of their presenting idiosycratic or inconsistent instructions.

Questionnaires as a Source of Experimental Artifacts:

Questionnaires used to assess the criterion variable are a possible threat to the internal and to the external validity (generality) of an experiment. Questionnaires may have both cueing and facilitating effects on the criterion variable, and their timing may affect the detection of significant effects.

Cueing effects of questionnaires

Questionnaires are a source of demand characteristics. The information contained in the questions presented to subjects provide cues as to the experimental hypothesis. This can influence subjects who are engaged in hypothesis guessing. The subjects' response to their guesses will depend upon their motivation to adopt a particular role in the experiment (see the discussion of subject roles, pp. 237-240). This cueing effect may be observed whether the questionnaire is first presented after the treatment (e.g., "'Aha, now I see why we got that movie'" (Campbell 1957, p. 280), or if it was administered prior to the treatment as well. Cueing effects may be more pronounced if the questionnaire precedes the treatment since they would then have additional information to use in interpreting and responding to the treatments as they were presented. In other words, the treatment may interact with the pretest. This effect may be somewhat dampened by consistency effects between the pre-treatment and post-treatment questionnaires.

Cueing effects may result in excess error variance if the questionnaires provide ambiguous cues. They may also result in falsely rejecting the null hypothesis if subjects accurately perceive the experimental hypothesis and are playing the "good subject" role (see page 238). Erroneous cues that are correlated with treatments may result in either falsely accepting or falsely rejecting the null hypothesis.

Facilitating effects of questionnaires:

Questionnaires may also facilitate the impact of the independent variable on the dependent variable. A pre-treatment questionnaire may create a mental set within which the treatment may create greater differences in the criterion variable than if no pre-treatment questionnaire had been administered. This synergism (interaction) may also be observed when only the post-treatment questionnaire is administered. "Note that this problem is conceptually different from the problems of demand characteristics.... We are not postulating that the subjects change their...attitudes because the experimenter wants them to, but only that no change would have taken place without the intrusion of a very obvious device for measuring the dependent variable..." (Carlsmith, Ellsworth and Aronson 1976, p. 198).

Because of the potential for both cueing and facilitating effects, pre-treatment questionnaires can rarely be justified. Their primary purpose is to use preteat scores as a covariate in the analysis or to examine change scores induced by the treatments.

The use of a control group which receives no treatment can provide the same information which could be obtained by examining change scores, without the interactive effects of a pretest.

Pretest scores may be used as covariates when the researcher thinks that the response to the treatment may depend on the pre-treatment level of the criterion variable. A pilot study could demonstrate if this approach should be used. Overall, pre-treatment questionnaires should be avoided whenever possible.

Normally, experimental treatments will not be confounded with the administration of pre-treatment questionnaires. However, in a complex

paradigm this confounding may accidentally occur (Baker and Churchill 1977 is an example).

Detecting and controlling cueing and facilitating effects of questionnaires:

Both the cueing and facilitating effects of pre-treatment questionnaires may be detected by comparing the performance of pretested to that of non-pretested groups on the post-treatment questionnaire. Both effects will be revealed statistically as a pretest-treatment interaction. Any interaction may be eliminated by eliminating the pre-treatment questionnaire. The cueing and facilitating effects of post-treatment questionnaires may only be detected by running experiments in which unobtrusive measures (Webb et al. 1966) are substituted for the post-treatment questionnaire. However, some effort should be made to ensure that the unobtrusive measures assess the same constructs as the questionnaire. The detection of facilitating and cueing effects of a posttreatment questionnaire dictates that experimental results observed in situations where the questionnaire was used may not be generalized to other settings.

The effects of questionnaire timing:

The effects of a treatment on a criterion variable may change over time. Therefore, a questionnaire administered after one period of time following presentation of the treatment may detect differences in the criterion variable across conditions whereas a questionnaire administered after a different time period may fail to detect any differences.

If this occurs, then the generality of the findings would be limited. "An experimental design which includes the measurement of the

dependent variable at several points in time will increase the ecological validity of the results" (Bracht and Glass 1968, p. 466).

Summary

This chapter has been concerned with methodological issues which influence a researcher's ability to draw causal inferences from his or her experimental results.

The requirements for valid causal inference were first presented. Following this, different sources of variance in the criterion variable which occur in experimental settings were identified. The implications of both systematic and random variance to drawing causal inferences were discussed along with various methods of controlling these types of variance.

The implications of sources of variance for the generality of experimental findings was also discussed along with methods for determining and specifying generalizations which may be made.

Finally, several persistent sources of experimental variance were examined: the experimenter, the subjects, the instructions and questionnaires which are used in experimental settings. Methods of controlling the systematic and random variance in the criterion variable which is attributable to these sources were discussed.

CHAPTER 5

HYPOTHESES AND METHODS

The present study was intended to examine the relationships among all parts of the elaboration model of message processing presented in Chapter one and three. This model is presented again for the readers convenience in Figure 8. The impact of all three types of antecedents on reported elaboration were to be examined as was the relationship between reported elaboration and post-message cognitive structure.

Hypothesized Relationships Between Elements of the Elaboration Model

Predicted Effects of the Strength of Prior Cognitive Structure on Reported Elaboration:

It was hypothesized that if an individual had based his prior cognitive structure (here operationalized as the individual's initial, i.e., pre-message, position) on information that was easily assailable he would tend to report a smaller proportion of negative to total responses (PROPNEG) to a discrepant message than would a person who had based his position upon information perceived to be unassailable. The person basing his initial position on relatively assailable information should be more unsure of his position and therefore less willing to defend it by responding negatively to a discrepant communication. The assailability of the information upon which the initial position was based was predicted to have the opposite effect on the proportion of reported positive to total responses (PROPPOS) to a discrepant communication.

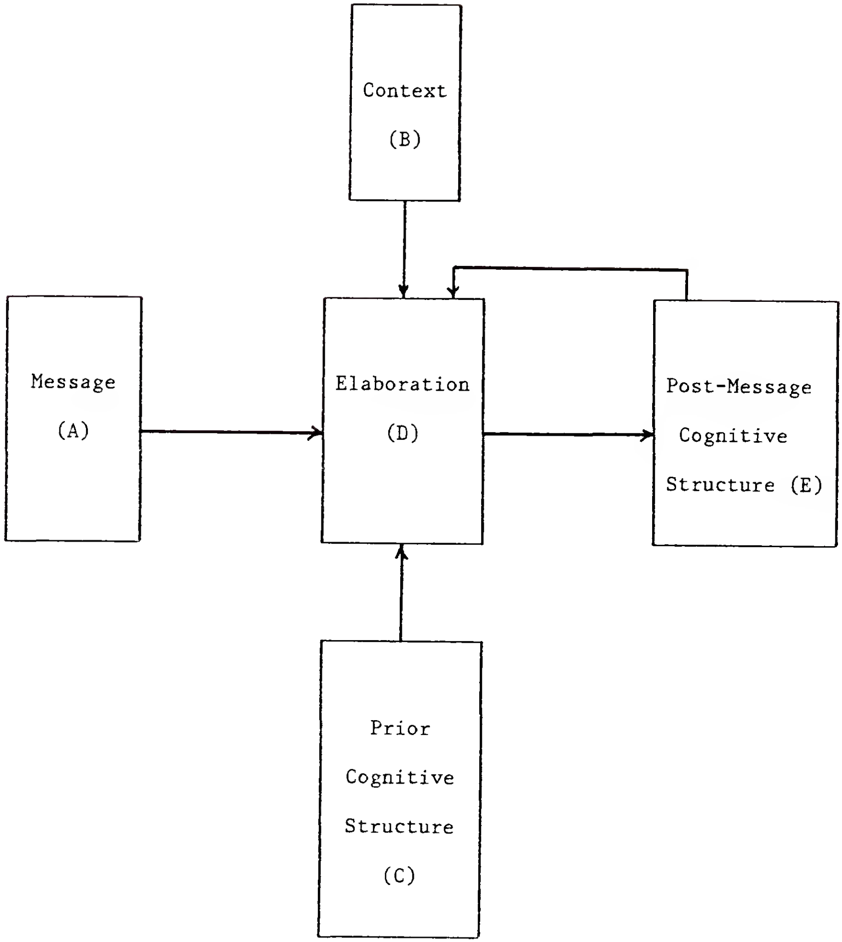


Figure 8. An Elaboration Model of Message Processing

Dissonance theory (Festinger 1957) would sometimes lead to predictions opposite to those presented here. Dissonance theory would lead to the prediction that, under some circumstances, a person who had knowingly accepted an assailable position would more strongly defend it than someone who had accepted an unassailable position. However, dissonance effects are generally observed when a person has made some commitment (especially public) to a position and when the position is tied in some way to the person's self-evaluation. However, subjects in the present study had made no commitment of any kind prior to receiving the discrepant communication. For that reason the dissonance predictions were not made in this study.

The Predicted Effect of the Strength of Prior Cognitive Structure on Post-message Cognitive Structure:

The literature discussed in Chapters 2 and 3 suggests that the resultant cognitive structure following a communicative episode will be a product of the coding which takes place through elaboration. It was therefore hypothesized that the impact of the strength of the prior cognitive structure on elaboration would also be reflected in measures of the post-message cognitive structure. Structural measures which would reflect acceptance of a discrepant communication should show greater acceptance for assailable than for unassailable position-anchoring information. Structural measures which indicate acceptance of the information upon which the initial position was based should indicate greater acceptance when the information is relatively unassailable than when it is easily assailable.

<u>The Predicted Relationship Between Reported Elaboration and Post-Message Cognitive Structure</u>:

Since the post-message cognitive structure is presumed to be a product of message elaboration, reported elaboration was predicted to be correlated with measures indicative of the post-message cognitive structure. This correlation should be observed if the measure of elaboration is a true report of elaboration and elaboration is an intervening process leading to a change in the cognitive structure. It should also be observed even if the <u>reported</u> elaboration is not a true measure of elaboration, since in this case it would be a <u>structural</u> measure and should be related to other structural measures.

As PROPNEG increases, structural measures indicating message acceptance should decrease. Alternatively, structural measures indicating acceptance of the background information should increase with an increase in PROPNEG directed at the discrepant communication.

PROPPOS was predicted to be correlated with the post-message structural measures in the opposite fashion to the correlations predicted for the PROPNEG. Thus, as PROPPOS increases, then structural measures indicating message acceptance should increase and structural measures indicating acceptance of the initial position should decrease.

<u>The Predicted Effect of Message-Source Credibility (a Contextual Variable) on Reported Elaboration and on Post-message Cognitive Structure</u>:

Respondents should show less tendency to defend their position when confronted with a discrepant communication when the source of that communication is high in credibility than when the source is low in credibility. Consequently, PROPNEG should be higher when source

credibility is low than when source credibility is high. The opposite should be observed for PROPPOS.

The effect of source credibility on post-message cognitive structure should correspond to the effects of source credibility on reported elaborative responses. Structural measures reflecting message acceptance should be higher for high source credibility than for low source credibility. Structural measures reflecting acceptance of the initial position should be lower for high source credibility than for low source credibility.

The Predicted Effect of Message Discrepancy on Reported Elaboration and Post-message Cognitive Structure:

The greater the discrepancy between the position expressed in the message and the position held by the message recipient, the greater should be the opposition expressed by the recipient. Therefore, it was predicted that PROPNEG would increase as the discrepancy between the position expressed in the message and the prior position of the recipient increased. Contrariwise, PROPPOS should decrease with increasing discrepancy.

According to social judgement theory (Sherif and Hovland 1961) acceptance of a proposed position should be negatively related to discrepancy. Acceptance should be greatest when discrepancy is small (i.e., when the proposed position falls within the recipient's latitude-of-acceptance about his position). When discrepancy is great, social judgement theory predicts negative persuasion (i.e., the "boomerang effect"). In this case the recipient's post-message position is further from that expressed in the message than was the recipient's pre-message position.

Social judgement theory would predict a curvilinear effect on change in position (PROPSHIF), with intermediate initial positions showing the greatest change. Intermediate positions are more ambiguous and are less stongly anchored than are extreme positions (see Kiesler, Collins and Miller 1969, Chapter 8).

Predicted Interactive Effects of Source Credibility and the Assailability of the Background Information:

If the recipient's initial position is strongly held (low assailability) changing from low to high source credibility should have less effect on reported elaboration and on structural measures than when assailability is high.

Likewise, a change in assailability should have a greater impact on reported elaboration and on structural measures when source credibility is low than when the source credibility is high.

Predicted Interactive Effects of Source Credibility and Discrepancy:

Kiesler, Collins and Miller (1969, pp. 292-297) discuss possible interactive effects of discrepancy and source credibility. They suggest that credibility along with other communicator characteristics "...can be converted into a good-bad evaluation. To the extent that the overall evaluation of the source is positive, assimilation tendencies should be increased; likewise, to the extent that they are unfavorable, contrast effects should be exaggerated (or superimposed)"(p. 292). Correspondingly, discrepancy was predicted to interact with source credibility.

Predicted Source Credibility X Background Information Assailability X Discrepancy Interaction on Background- and Message-Specific Measures of Acceptance:

Two criterion measures evaluated the likelihood that factors presented as central to the "test" (anionic surfactant content and grain

size) determine a detergent's cleaning ability (LIKANSUR and LIKGRSIZ). Similarly, three factors were presented as alternative determinants of cleaning ability in the message (water temperature, sudsiness, and phosphorus content). Three criterion measures assessed the acceptance of these factors by measuring the perceived likelihhod that these factors determine cleaning ability (LIKWATMP, LIKSUDS, and LIKPHOS).

It was predicted that discrepancy would interact with source credibility and background information assailability on these criterion measures. To account for the discrepancy between the position presented in the background information and that claimed in the message, the subject could attribute the difference to either a lack of efficacy of the factors utilized in the "test" of those claimed in the message. It was predicted that increasing discrepancy would lead to decreases in LIKANSUR and LIKGRSIZ and increases in LIKWATMP, LIKSUDS and LIKEPHOS when source credibility was high and background information assailability was high. The opposite effect of discrepancy was predicted when source credibility and background information assailability were low. (Unfortunately, because of the inadequacy of the manipulations for source credibility and background information assailability this interaction could not be tested. The cell means for LIKANSUR, LIKGRSIZ, LIKWATMP, LIKSUDS and LIKPHOS are presented in Table 15 in the Appendix.)

Plan of the Experiment

To test the relationships hypothesized in this study it was decided that there should be two levels of assailability of the background information and two levels of message source credibility. This would

enable the experiment to detect any interaction between the two variables and to test the impact of a variable not previously examined--the assailability of the information upon which a threatened position is based.

Three levels of discrepancy were used to determine if the social judgement phenomenon would be exhibited across a range of discrepancy. Message and no message conditions were used so that the effect of each discrepancy condition could be assessed by comparing it with a no message condition. The combinations of these variables resulted in a nested, factorial design.

Method:

Experimental design:

The design of the experiment consisted of a nested, factorial design as is illustrated in Figure 9. The message-no message factor was crossed with the assailability of the background information. Discrepancy and message-source credibility were crossed and nested within the message condition of the message-no message factor.

Setting:

To better understand the experiment, it is important to visualize the setting in which the study was performed. The subjects were seated by two's side-by-side at eight-foot tables. Partitions (wooden frames covered with unbleached muslin material) divided each table at the middle and thereby separated the two subjects' stations. Partitions also enclosed the table at each end and across the back. The partitions extended out from the side of the table on which the subjects were seated. Each station was arranged so that no subject could see what the others were doing. This was done to minimize social interaction effects.

Figure 9. Experimental Design.

	MESSAGE						NO MESSAGE
	Low Discrepancy		Moderate Discrepancy		High Discrepancy		
	high CC	Low CC	high CC	low CC	high CC	low CC	
Low Assailability	1	2	3	4	5	6	7
High Assailability	8	9	10	11	12	13	14

Subjects:

Subjects for the sessions were solicited by an independent contractor (Central Florida Research, Gainesville, Florida). Subjects were obtained through intact groups in the Gainesville area which were able to earn money by their members' participation in the sessions. The groups received five dollars for each subject's participation. All subjects participating were housewives.

Procedure:

Upons a subject's arrival, she was randomly assigned to a station which determined the condition she would receive in the experiment. After the experimenter introduced himself and his four assistants the study was described as one being conducted under the auspices of the Center for Consumer Research. Subjects were told that the study was free of the sponsorship of any one detergent manufacturer. Subjects then completed informed consent forms.

After subjects had been assigned to stations, but prior to the present study, they participated in another experiment. In this prior experiment they were required to read various background material regarding a potential detergent cleaning ability labeling program. Under this program, detergents would be assigned a cleaning ability rating based on a washing test. In addition to the background reading, subjects also engaged in a simulated shopping trip which involved the selection of one box of a branded detergent from among five brands which were present at their stations. Subjects also responded to various paper and pencil measures regarding detergents.

Participation in the first experiment had two positive effects. First, it served to arouse subject involvement in the topic of detergent

cleaning ability ratings. Second, since the proposed labeling program had been based on washing test comparisons, the first experiment also served to increase experimental realism by providing the setting and the justification for the presentation of the formula-based test in the background reading for the present study. There are also two possible negative effects of participation in the first experiment. First, subjects could be fatigued after completing the first experiment. This could serve to reduce subject involvement and cooperation. Second, subjects acquired differential experience across conditions of the first experiment. To avoid confounding of conditions of the first experiment with those of the experiment presently being examined, the conditions of the first experiment were balanced across the conditions of the second (present) experiment. However, the effects of the first experiment still remain as a possible source of error variance.

In the present study, subjects were first given background reading material regarding a "government-sponsored" physical and chemical evaluation scheme for detergents as an alternative to the washing test proposed in experiment I. After they had finished reading the background material (see the appendix), they were given a letter purported to be from a manufacturer responding to the government-sponsored scheme. When subjects had finished reading the letter, they were given a thought-listing questionnaire in which they could list thoughts that they had in response to the letter.

Following the thought-listing questionnaire, subjects were given a final questionnaire in which they were asked to respond to various scales. Following that questionnaire the subjects were debriefed and dismissed.

Background reading material:

The background reading material presented a formula-based "test" as a way of predicting the cleaning ability rating which a detergent would receive on a washing test. This formula-based "test" was to be based on the anionic surfactant content and grain size of the detergent. The predicted cleaning ability of each of three bogus detergent brands was projected on the basis of this proposed system. (To see how the attributes combined, see the background reading material in the Appendix.) These projected levels comprised the initial positions held by the subjects.

The background reading effected the manipulation of one of the predictor variables--the assailability of the background information. Here the intent was to vary how strongly held the threatened position would be (a prior cognitive structure variable). The variable closest to this one from previous studies is commitment (Maccoby and Roberts 1972; Roberts and Maccoby 1973; and Gardner 1966, 1970). When the initial position of the subjects is based on an unassailable information, it should be more resistant to change induced by the message. The source of the initial position was the same in both conditions. However, the "test" which was used to derive the cleaning ability ratings varied. In the high assailability condition the testing procedure was characterized as one "...selected five months ago by the government detergent program staff from among the few formula-based tests that are available." In the low assailability condition, the procedure was positioned as one that "...was developed five years ago. Its procedures are well established and have been approved by the Tests and Procedures Committee of the American Chemical Society."

Subjects were asked to turn the background reading booklet over to signal that they had finished reading.

Letter from the manufacturer:

When all subjects had finished reading the background material, they were given a mock letter from the manufacturer of one of the three brands whose rating had been predicted using the formula-based system in the background reading. The letters were identical for each manufacturer with the exception of the brand designation. The letters varied, depending on which credibility condition they represented--either being attributed to a high credibility or a low credibility source.

In the high credibility condition, subjects received a letter which began: "We in the Product-Testing Department would like to tell you some of our opinions...." In the second paragraph the letter continues: "As Vice-President in charge of Product Testing, I feel confident in assuring you..." and "we have already gone ahead with the development of an introductory advertising campaign proclaiming the excellence of our product." The letter is signed "David Richards, Ph.D. Chem., Vice-President, Product-Testing Department."

In the low-credibility condition, "Advertising Department" is substituted for "Product-Testing Department" in the first paragraph. In the second paragraph, "As Vice-President in charge of Advertising" replaced "As Vice-President in charge of Product-Testing," and "we have already gone ahead with the adaption of our present formula and preparations for production" replaced the statement about the introductory advertising program. This letter was signed by "David Richards, Vice-President, Advertising Department."

The letter also completed the discrepancy manipulation by claiming a rating for the letter's sponsor which was discrepant with the position indicated by the formula-based test. (See the appendix for a sample letter.) Subjects were instructed that they could refer to the background reading as they read the letter.

Discrepancy could have been manipulated by making differential claims in different message conditions (varying extremity) and relying on the random assignment of subjects to roughly equalize the initial positions held by Ss across conditions. This is the approach that was taken by Brock (1967). There are two disadvantages of this approach. First is the lack of control over the initial position held by the subjects. Second is the possibility that different claims vary in their inherent believability because of the level of the position proffered, rather than because of discrepancy.

Another approach would be to experimentally establish one position on a topic as the one held by all subjects. This should be one for which they could have had no prior position. These subjects would, then, be presented with messages expressing different positions in the various discrepancy conditions. Again, there is the problem of possible differential persuasiveness resulting from the <u>extremity</u> of the advocated position, rather than from its distance from the position held by the subjects (i.e., extreme claims may be inherently less believable).

One way of avoiding this problem whould be to cross these two variables in a factorial design such that each message position occurs with each initial position. Despite the attractiveness of this approach, it was not utilized in this study because of the burgeoning

number of conditions that would result, given the levels of the other predictor variables and because it would be unrealistic in the setting for a claim to be made which was lower than the initial position.

The approach that was used in this study was one in which three positions were experimentally established as part of all subjects' prior (to message) cognitive structures. These positions represented the predicted cleaning ability of the three bogus-brand laundry detergents, as predicted by the formula-based "test" and were presented as part of the background reading, both numerically and visually with an arrow pointing to the predicted cleaning ability of each detergent. Discrepancy conditions were effected through messages which all claimed a cleaning ability of Excellent/Grade AA (superior to all of the positions established in the prior structure), but varied as to the brand making the claim. (All Ss had been exposed to the labeling proposal in Experiment 1). In effect, discrepancy results from an interaction between the message and the subjects' prior cognitive structure.

This procedure eliminates the problem of differential persuasibility due to the level of the claim. However, the three initial positions may have differed in their inherent assailability (another level effect)--independently of the position presented in the message. Unfortunately, either assailability of the initial position or the persuasibility of the message due to level must be confounded with discrepancy if the claim is to exceed the initial position. Since subvocal responses to the message were to be elicited, it was decided to leave the claims free of this confounding. The differential assailability of the three initial positions thus remains as an alternative explanation for any effects of discrepancy observed in this study.

When they had finished reading the letter (a maximum of three minutes was allowed), they were to raise their hands, whereupon one of the experimenter's assistants would collect the background reading and the letter and give them the Thought-listing Questionnaire. (See the appendix.)

Thought-listing questionnaire:

Thought listing was utilized to gather reported elaborations. (Also recall the discussion in Chapter 3 of the potential of thought-listing as a measure of persuasion which is relatively non-reactive compared to more direct measures of persuasion.)

The time elapsed between reading the letter and beginning the thought-listing task included the delay for presentation of the instructions, as well as a waiting period for those readers who finished early. Unfortunately, this variable waiting period was unavoidable, given the group nature of the experiment's administration and represents a source of error variance.

The thought-listing procedure that was used was adapted from those used by Petty, Wells, and Brock (1976) and by Wright (1973a). Subjects were told via written instructions, which were also read aloud by the experimenter, that "We are now interested in whether or not thoughts related to the communication might have occurred to you while you were reading the manufacturer's reply. On the next page you can record any of these thoughts that you can recall." This differed from the Petty, Wells, and Brock (1976) approach in that it referred specifically to thoughts that may have occurred during the presentation of the message (in this case, reading). However, it is the same approach as that used by Wright (1973a).

The instructions were phrased to avoid the reactive listing of thoughts by subjects to impress the experimenter by stating: "If you honestly can't recall thinking about what was said, that's fine, simply write NONE on the first line." The instructions used here were also similar to Wright's approach, in that they instructed subjects to list any thoughts and did not suggest possible polarity of the thoughts as did the instructions used by Cullen (reported in Greenwald 1968); Calder, Insko, and Yandell (1974); Petty, Wells, and Brock (1976); and Cialdini et al. (1976). (See Chapter 3). (For the complete text of the instructions see the appendix).

At the bottom of the page, subjects were informed that they would have 2½ minutes to write their thoughts. The next page had the heading: "Thoughts that occurred to you as you were reading the manufacturer's reply. Please put only one thought per box." On that page were nine boxes formed by ten hortizontal lines, one inch apart and joined by a vertical line 3/4 inch from the left margin.

When the listing time had expired, subjects were asked to turn to the last page of the thought-listing questionnaire. Again, they were asked to follow along as the experimenter read the instructions aloud. They were instructed to turn back to the page on which they had listed their thoughts and write either a plus or a minus next to each thought. "If what you wrote seems to be in favor of what the manufacturer had to say, then write a plus (+) in the left-hand margin next to that thought. If what you wrote seems to be opposed to what the manufacturer had to say, then write a minus (-) in the margin next to that thought." This measure was not analyzed because subjects disagreed with the judges in many cases as to both the polarity and the separation of thoughts.

Since the judges were working from operational definitions that could be applied across subjects, the judges' coding of thoughts was used for analysis.

The thoughts were coded acording to a priori operational definitions (see appendix) as counterarguments, support arguments, source derogations, or source enhancements. The first three categories had been used in previous research (Wright 1973a). The fourth, source enhancement, was suggested by Professor Albert R. Wildt during the presentation of a proposal for this study at the University of Florida.

Three judges who were unaware of the conditions which the Ss were in coded the statements independently. Their codings were compared and final assignments were made to categories. If all three judges or if two of the three agreed on a category then that category was used. If all three disagreed, the statement was assigned to the Neutral category.

When the listing was completed this questionnaire was collected and the Preference Questionnaire was distributed.

Preference questionnaire:

As the subjects followed along, the experimenter read the instructions on the first page of the Preference Questionnaire. These instructions (see appendix) explain how the Ss should use the preference scales on the second page of the questionnaire. When the instructions were concluded, subjects were asked to turn the page and to look at the first pair to be rated. The instructions were repeated for this pair and for the second one (a pretest had found that subjects had difficulty in applying the instructions without this additional prompting).

The paired-comparison preference measure which was used was adapted from Bechtel (1976, p. 78). The order of presentation of the pairs and

the left-right ordering of brands within pairs was counterbalanced across subjects to eliminate possible order effects.

An individual's utility for each brand may be derived from his preference scores. This is done by arraying the preference scores for each pair as a matrix which presented all of the possible pairs (P_{MN} = how much S prefers M to N; $P_{NM} = -P_{MN}$; $P_{MM} = P_{NN} = P_{LL} = 0$):

$$\begin{matrix} P_{MM} & PMN & P_{ML} \\ P_{NM} & P_{NN} & P_{NL} \\ P_{LM} & P_{LN} & P_{LL} \end{matrix}$$

Averaging across the rows of this matrix yields the individual's <u>utilities</u> for the alternatives (Bechtel 1976, p. 18):

$$U_M = P_{MM} + P_{MN} + P_{ML}$$
$$U_N = P_{NM} + P_{NN} + P_{NL}$$
$$U_L = P_{LM} + P_{LN} + P_{LL}$$

The program which derives these utilities subtracts the mean of U_M, U_N, and U_L from each utility to "column center" them. Thus,for each individual, these utilies summed to zero. It was initially intended that the derived utility measures would be compared with other structural measures through correlational analysis. However, later discussion with Professor Bechtel indicated that cross-sectional correlations involving derived utility measures are inappropriate since these measures are not directly comparable on an inter-subject basis. Therefore the planned analysis was dropped.

When the subjects had completed the preference rating, they were asked to turn the page to the Formula-Based Rating System Questionnaire (see appendix).

Formula-based rating system questionnaire: The two questions on the first page were measures of post-message cognitive structure. The first assessed the likelihood "...that a cleaning ability rating of Excellent/Grade AA is appropriate for each of the ...brands" (LIKEXM, LIKEXN, LIKEXL). This was the cleaning ability claimed in each manufacturer's message. For question 1 the experimenter gave the following instructions: "Look at question number 1. How likely do you think it is that a cleaning ability rating of Excellent/ Grade AA is appropriate for brand M? Look at the line for brand M. Find the number that expresses how likely you think it is that a cleaning ability rating of Excellent/Grade AA is appropriate for brand M and circle that number. Now do the same thing for brands N and L".

The second question asked subjects to rate the detergents utilizing a scale identical to the one by which the initial cleaning abilities of the detergents were presented. The positions on the scale which subjects thought represented the cleaning ability of the detergents were taken as a measure of their post-message cognitive structure (RATM, RATN, RATL). A pretest had indicated that subjects had some difficulty with this question. As a result additional prompting was given by the experimenter.

When the subjects had completed question 1, they were told: "Take a minute to read question 2 and then I'll go over it with you." When they had had a chance to read over the question, the experimenter continued: "Think about brand M. What do you think is the actual cleaning

ability rating that this brand would be assigned following the washing tests? Find that point on the line that you think brand M would be assigned following the washing tests. Start above that point on the line and draw an arrow straight down to that point. Now write an M above the arrow." These instructions were repeated for brand N and subjects were asked to continue in the same fashion for brand L.

To code the response to this question, the position of the arrows was determined to the nearest 1/10 of a unit using a ruler especially constructed for that purpose. The proportion of the distance from the initial position to the Excellent/Grade AA category which the subjects shifted the position of the arrow (PROPSHIF) was calculated for each brand (PROPDISM, PROPDISN, PROPDISL). This proportion was used as an indicator of a shift in cognitive structure. A proportion was used since the distance from the initial position to the Excellent/Grade AA category was confounded with discrepancy. If distance rather than a proportion had been used the most highly rated brand could have shown little absolute movement, yet have shifted to the cleaning ability rating claimed in the message.

When subjects had completed the rating task, they were asked to turn the page and complete the rest of the questionnaire, which consisted of straightforward multichotomous questions. The first four items were used to assess the subjects' responses to the manufacturer's reply. These items were ad hoc scales which elicited how <u>believable</u> the manufacturer's claim was (COMBEL), how <u>trustworthy</u> (COMTRUST) and <u>justified</u> (COMJUST) the reply was, and how <u>likely</u> it was to have been based on <u>expertise</u> (COMEXPRT). These scales were constructed by using variations of a key word from each statement as the response categories. Six

categories were used for each scale. For example: Very Believable (scale value of 1), Believable, Somewhat Believable, Somewhat Unbelievable, Unbelievable, Very Unbelievable (scale value of 6). The key words used for the other three scales were Trustworthy, Justified, and Likely.

These questions were intended as indicators of cognitive structure and as post-hoc manipulation checks of the manufacturer's credibility manipulation. Of course, as post-hoc measures, they may be influenced by persuasion and may reflect consistency with earlier measures in addition to revealing the impact of the manipulation.

The next four ad hoc scales were utilized to assess the subjects' evaluation of the formula-based rating system. They were constructed in the same fashion as those used to evaluate the manufacturer's reply. Again, six categories were used with the positive extreme assigned a 1 and the negative extreme a 6. The key words used were accurate (TESTACC), representative (TESTREP), appropriate (TESTAPP), and trustworthy (TESTRUST). In the interest of space in constructing the alternate versions of the questionnaire, only the last three scales were included for the treatment groups.

The last five scales were used to assess acceptance of the formula-based rating system and the manufacturer's reply. Using the keyword, Likely, ad hoc scales were constructed to measure how likely Ss thought it was that certain specific factors determined detergent cleaning ability. The factors examined were: a) anionic surfactant content of a detergent (LIKANSUR); b) detergent grain size (LIKGRSIZ); c) water

temperature (LKWATMP); d) sudsiness (LIKSUDS); and e) phosphorus content (LIKPHOS)*.

Two of the factors (anionic surfactant content and grain size) formed the basis of the formula-based rating system. The other three were mentioned in the manufacturer's reply as moderating factors which limit the usefulness of simple chemical and physical description in predicting a detergent's cleaning ability.

Subjects were asked to raise their hands to indicate when they had completed the questionnaire so that it could be collected. When all questionnaires had been collected, subjects were debriefed and dismissed (see appendix for the complet text of the debriefing).

Analysis:

Effects of predictor variables on the criterion measures:

All of the hypothesized relationships in this study were directional. Therefore, t-tests were utilized to test a priori contrasts to examine the effects of discrepancy on the criterion variables (recall that the manipulations of communicator credibility and test assail ability were unsuccessful). An F-ratio could have been used to test these hypotheses. However, since the numerator is squared in the F-ratio, the directionality of the comparison is lost.

The contrasts were tested using the General Linear Model (GLM) procedure of the 1979 version of SAS (Statistical Analysis System (SAS Institute, 1979)). When testing contrasts, the output of this procedure is an F statistic and its associated probably. To obtain the t-value, the square root of the F value was taken. "An $\underline{F}$ ratio for 1 and v

*Note: The scale values for LIKEXM, LIKEXN, LIKEXL, COMBEL, COMTRUST, COMJUST, COMEXPRT, TESTACC, TESTAPP, TESTREP, TESTRUST, LIKANSUR, LIKGRSIZ, LIKWATMP, LIKSUDS andLIKPHOS are all inversely related to the expressed beliefs.

degrees of freedom is equivalent to t^2 and ν degrees of freedom" (Kirk 1968, p. 81). Since the contrasts were directional, one-tailed t-tests were appropriate. Therefore, to obtain the probability of t under the null hypothesis, the probability associated with the F values was divided by two (since the F-test is non-directional).

Relationships between reported elaboration and structural variables:

The relationships between reported elaboration and structural variables was examined through the correlational analysis proceedure of the 1979 version of SAS (SAS Institute 1979). Since the predicted correlations were directional, the probability associated with the observed correlation coefficient was obtained by dividing the probability from the proceedure's output by two.

Internal analysis:

Because of the failure of the source credibility and information assailability manipulations, the effects of these variables on the various criterion measures could not be tested. However, Carlsmith, Ellsworth and Aronson have suggested the usefulness of an internal analysis in such situations (1976, p. 37). An internal analysis examines the relationship between the manipulation check measures (as a surrogate of the predictor variables) and the criterion measures. This approach provides useful insights regarding the possible underlying relationships, but, being correlational, cannot support causal statements about these relationships.

An internal analysis was performed using the manipulation check measures: COMBEL, COMTRUST, COMJUST, COMEXPRT, TESTAPP, TESTREP, and TESTRUST. The correlations of these measures with the criterion

measures were compared with directional predictions in the same manner as described above. The results of this internal analysis and other analyses are presented and discussed in the next chapter.

CHAPTER 6

RESULTS AND CONCLUSIONS

This chapter will present the results and conclusions from this study, as conducted. Because the manipulations regarding source credibility and background information assailability were unsuccessful, the results associated with the effects of those manipulations on criterion variables will not be presented (see Tables 16 and 17 in the appendix for manipulation check measures).

The effects of discrepancy on reported elaboration and on structural measures will be reported. A second set of relationships presented will be those between reported elaboration and structural measures. The final set of relationships presented comprise an internal analysis (Carlsmith, Ellsworth and Aronson 1976, p. 37). This internal analysis examines the relationships of the measured indicants of source credibility and information assailability with reported elaboration and with structural variables. The final part of this chapter will present suggestions for future research regarding the relationships examined in the present study as well as other relationships suggested by the elaboration model of message processing.

Effects of Discrepancy

This section presents the findings of the a priori contrasts that were preformed to examine the effects of discrepancy. The a priori contrasts were evaluated by the use of t-tests which utilized the MSE (mean square error) from the general linear model (GLM) solution of the 1979 version of SAS (Statistical Analysis System (SAS Institute Inc.,

1979)). Since the hypotheses tested by contrasts were directional, t-tests were used instead of F-tests. F-tests are similar to t-tests for contrasts except that the ratio is squared for F-tests. The squaring of the numerator results in the loss of directional information (Kirk 1968, pp. 74 & 81).

Effects of Discrepancy on Reported Elaboration:

Table 9 presents the cell means for PROPNEG and PROPPOS. The effect of the three predictor variables were tested using a priori contrasts.

The discrepancy treatment means for PROPNEG were: low discrepancy (Brand M) = 0.292, moderate discrepancy (Brand N) = 0.407, and high discrepancy (Brand L) = 0.456. Contrasts showed that the null hypothesis could not be rejected throughout the range of discrepancy present in this study. The contrast which compared the low and moderate conditions was significant ($t_{177} = 1.67$, $p = 0.048$). However, the comparison of the moderate and high conditions was non-significant ($t_{177} = 0.75$, $p = 0.22$).

PROPPOS was unaffected by discrepancy throughout the discrepancy range in the present study. The t-ratio for the low-moderate and moderate-high contrasts were $t_{177} = 0.67$ ($p = 0.25$) and $t_{177} = 0.73$ ($p = 0.23$), respectively. Thus, the null hypothesis that the treatment means of 0.340 (low), 0.294 (moderate), and 0.255 (high) are equal was not rejected.

Effects of Discrepancy on Detergent Ratings:

Since discrepancy was confounded with the initial brand rating, the effect of discrepancy was tested by comparing the rating of the brand from whom the Ss received the letter with the rating of that same brand

Table 9. Cell Means for the Proportions of All Protocol Statements Positive and Negative (CC = communicator credibility).

		MESSAGE						NO MESSAGE
		Low Discrepancy		Moderate Discrepancy		High Discrepancy		
		high CC	Low CC	high CC	low CC	high CC	low CC	
	Criterion							
Low Assail-ability	PROPNEG	0.292	0.200	0.463	0.227	0.471	0.531	--
	PROPPOS	0.398	0.406	0.256	0.456	0.282	0.147	--
		n = 15	n = 14	n = 16	n = 15	n = 15	n = 15	
		1	2	3	4	5	6	7
High Assail-ability	PROPNEG	0.366	0.291	0.477	0.449	0.403	0.426	--
	PROPPOS	0.306	0.286	0.343	0.153	0.364	0.230	--
		n = 18	n = 16	n = 15	n = 18	n = 15	n = 17	
		8	9	10	11	12	13	14

by subjects who received no letter. Brand rating means are presented in Table 10.

The means of RATM for those who received the letter from Brand M (7.05) and those who received no letter (7.18) differed in the opposite direction to that which was predicted. The null hypothesis was, therefore, not rejected.

The means of RATN for the N-message condition (5.75) and for the no-message condition (5.37) were not significantly different by a two-tailed t-test ($t_{210} = 1.40$, $p = 0.16$). Again, the null hypothesis was not rejected.

The mean RATL for the L-message group (3.89) was higher than that for the no-message group (3.15). This was opposite to the direction predicted, so the null hypothesis was not rejected. This difference was significant in a two-tailed t-test ($t_{209} = 2.24$, $p = 0.03$).

The predicted assimilation and contrast effects of discrepancy on brand ratings were not observed. The only significant finding was an unpredicted one--an increase in RATL for those Ss receiving the L-message over RATL for Ss receiving no message.

Effects of Discrepancy on the Shift in Detergent Rating:

The effect of discrepancy on each proportional shift was tested by comparing the shift for the same brand which had sponsored the message with the shift for that brand by the no-message condition. Proportional shift means are presented in Table 11.

The contrast for PROPDISM (M-message low discrepancy vs no-message) found that the message mean (0.88) and the no-message mean (0.38) were not significantly different ($t_{215} = 0.54$, $p = 0.30$). Thus, there was a failure to reject the null hypothesis.

Table 10. Cell Means for Detergent Ratings (CC = communicator credibility).

		MESSAGE						NO MESSAGE
		Low Discrepancy		Moderate Discrepancy		High Discrepancy		
	Criterion	high CC	Low CC	high CC	low CC	high CC	low CC	
Low Assail-ability	RATM	6.99	6.92	7.54	7.51	7.02	7.57	6.97
	RATN	5.55	4.92	5.33	6.32	5.84	5.92	5.28
	RATL	3.34	4.08	3.91	4.67	3.48	3.45	3.02
		n = 14	n = 13	n = 16	n = 15	n = 13	n = 15	n = 19
		1	2	3	4	5	6	7
High Assail-ability	RATM	7.41	6.80	7.72	7.53	7.38	7.26	7.38
	RATN	5.73	5.96	5.36 (14)*	5.96	6.04	5.29	5.45
	RATL	3.86	4.12	3.75	3.23	4.07	4.45	3.27
		n = 18	n = 16	n = 13	n = 18	n = 15	n = 17	n = 21
		8	9	10	11	12	13	14

*All means within a cell are based upon the same n except as shown in parentheses.

Table 11. Cell Means for the Proportion of the Distance to the Excellent/Grade AA Category which the Brand was Shifted. (CC = communicator credibility)

		MESSAGE						NO MESSAGE
		Low Discrepancy		Moderate Discrepancy		High Discrepancy		
		high CC	Low CC	high CC	low CC	high CC	low CC	
	Criterion							
Low Assail-ability	PROPDISM PROPDISN PROPDISL	1.57 0.58 0.31	1.57 0.37 0.48	1.09 0.13 0.20	1.03 0.53 0.37	3.24 1.04 0.57	1.13 0.37 0.10	-0.04 0.11 0.005
		n = 15	n = 14	n = 16	n = 15	n = 15	n = 15	n = 19
		1	2	3	4	5	6	7
High Assail-ability	PROPDISM PROPDISN PROPDISL	0.82 0.29 0.19	-0.40 0.28 0.25	4.44 0.51 0.62	1.06 0.38 0.05	2.31 0.42 0.24	0.52 0.12 0.32	0.75 0.18 0.06
		n = 18	n = 16	n = 15	n = 18	n = 15	n = 17	n = 21
		8	9	10	11	12	13	14

The contrast for PROPDISN (moderate discrepancy) also failed to reject the null hypothesis. The means for the message (0.38) and the no-message (0.14) conditions were not significantly different (two-tailed $t_{215} = 1.27$, $p = 0.21$).

Since the comparison for PROPDISL (high discrepancy) showed the means varying in the opposite direction to that hypothesized, the null hypothesis was not rejected. The mean for the L-message condition (0.31) and that for the no-message condition (0.03) were significantly different by a two-tailed t-test ($t_{215} = 2.12$, $p = 0.035$).

Overall, the effect of discrepancy on PROPSHIF did not support the hypotheses. There was only one significant contrast and that was in the unexpected direction. PROPDISL was significantly increased over the no-message condition by giving Ss a message claiming that Brand L deserved a cleaning ability rating of Excellant/Grade AA. Neither of the other brands was similarly affected.

Effects of Discrepancy on the Likelihood of an Excellent/Grade AA Rating's Appropriateness for a Brand:

All but one of the contrasts to examine the effects of discrepancy, on LIKEXM, LIKEXN, and LIKEXL were tested by the use of one-tailed t-tests. The exception was the discrepancy comparison for LIKEXN, which did not involve a directional prediction. For each of these measures, the smaller the mean, the greater the perceived likelihood of excellence. The cell means for LIKEXM, LIKEXN, and LIKEXL are presented in Table 12.

The effects of discrepancy on LIKEXM, LIKEXN, and LIKEXL were tested by using contrasts which compared the means of each brand's

Table 12. Cell Means for the Likelihood of a Brand's Excellence (CC = communicator credibility).

		MESSAGE						NO MESSAGE
		Low Discrepancy		Moderate Discrepancy		High Discrepancy		
	Criterion	high CC	Low CC	high CC	low CC	high CC	low CC	
Low Assail-ability	LIKEXM	2.33	2.14	2.44	2.40	2.33	2.33 (15)*	2.16
	LIKEXN	3.00 (14)*	3.36	3.69	2.73	2.73	3.07	2.79
	LIKEXL	4.07	3.50	4.00	3.33	3.73	4.43	4.16
		n = 15	n = 14	n = 16	n = 15	n = 15	n = 14	n = 19
		1	2	3	4	5	6	7
High Assail-ability	LIKEXM	2.00	2.25	1.64	2.00 (18)*	2.13	2.12	2.48
	LIKEXN	2.56	2.50	2.87 (15)*	2.76	2.60	3.12	3.48
	LIKEXL	3.18 (17)*	3.00	3.64	3.59	3.93	3.59	4.71
		n = 18	n = 16	n = 14	n = 17	n = 15	n = 17	n = 21
		8	9	10	11	12	13	14

*All means within a cell are based upon the same n except as shown in parentheses.

likelihood of excellence from the group which received a letter from that same brand's manufacturer with the mean of that brand's likelihood of excellence from the no-message condition.

The mean of LIKEXM for the M-message condition (2.17) was less than that for the no-message condition (2.33). Although this difference indicates a greater perceived likelihood of excellence for the M-message condition, this difference was not significant in a one-tailed t-test ($t_{214} = 0.62$, $p = 0.27$). Therefore, the null hypothesis was not rejected.

LIKEXN was not significantly different between the N-message (3.02) and the no-message (3.15) conditions (two-tailed $t_{212} = 0.56$, $p = 0.58$). This was as predicted. The null hypothesis was not rejected.

LIKEXL was greater in the no-message (4.45) condition than in the L-message condition (3.90). This was opposite to the direction predicted. This difference approached significance in a two-tailed t-test ($t_{211} = 1.82$, $p = 0.07$).

The assimilation-contrast effects which were predicted for discrepancy on likelihood of excellence were not observed. Even though the expected lack of difference for LIKEXN was observed, the overall pattern of findings did not support the predictions. Only one contrast approached significance. That was the difference for LIKEXL between the L-message condition and the no-message condition, which was in the unpredicted direction. This suggests that the most discrepant message within the range represented within this study may have the greatest chance of significantly persuading the recipients. The lack of a significant finding for the most highly rated brand (M) may reflect a ceiling effect.

Summary of Discrepancy Effects:

Discrepancy was found to partially affect reported elaboration in the predicted manner. Increasing discrepancy from low to moderate increased PROPNEG. PROPPOS was unaffected.

These findings are not strong support for the predicted relationship. However, this study utilized an alternative way of operationalizing discrepancy compared with that used by Brock (1967). The convergence of the present findings relative to post-message protocol measures with those found by Brock (1967) increases the likelihood of the underlying relationship between discrepancy and elaborative responses.

While social-judgement (assimilation and contrast) effects were predicted for the effect of discrepancy on detergent evaluation measures, they were not observed. It was only for the most discrepant claim that significant effects were observed. The detergent rating and the proportion of the distance to the Excellent/Grade AA category that that rating was shifted relative to the means for the no-message condition was significantly higher for the evaluation of the high discrepancy brand (Brand L) by those receiving the highly discrepant message. Thus, the more discrepant the message, the more effective it was in increasing the perceived evaluation of the referenced detergent. The effect for the perceived likelihood of excellence of Brand L, while not significant, reflected the same trend as the rating and proportional shift.

At least for the range of discrepancy present in this study, the assimilation and contrast effects did not occur.

Reported Elaboration--Post Cognitive Structure Relationships

The relationships between reported elaboration and other criterion measures was examined through correlational analysis. The results are summarized in Table 13. For ease of interpretation, all observed correlations with the variables are presented as the correlations between the underlying variables, regardless of the coding for those variables. (Lower scale values for some variables indicated a greater level of the dimension in question, i.e., greater likelihood, believability, appropriateness, etc.)

Observed Relationships Between Reported Elaboration and Detergent Evaluations:

The relationship between reported elaboration and detergent evaluations was examined such that the brand being evaluated was the same as that of the source of the communication toward which the reported responses were directed.

Observed relationships between reported elaboration and likelihood of a brands's excellence:

The likelihood of a brand's excellence was a specific issue of contention in the message. All of the observed correlations were in the predicted direction.

It appears that PROPNEG was a better predictor of the likelihood of a brand's excellence than PROPPOS, as the former was significantly correlated to both LIKEXM and LIKEXL. PROPPOS was significantly related to LIKEXL in the predicted direction. Thus, as PROPNEG increased and PROPPOS decreased, LIKEXL decreased. As PROPNEG increased, LIKEXM decreased.

Table 13. Correlations of Reported Elaboration with Structural Measures.*

	PROPNEG			PROPPOS		
Criterion	r	P(r) H_0	n	r	P(r) H_0	n
LIKEXM brand = M	-.38	.001	63	.14	.13	63
LIKEXN brand = N	-.07	.30	63	.10	.22	63
LIKEXL brand = L	-.28	.01	61	.37	.002	61
RATM brand = M	-.14	.13	61	.02	.44	61
RATN brand = N	-.09	.23	63	-.02	.87**	63
RATL brand = L	-.20	.06	60	.37	.002	60
PROPDISM brand = M	-.12	.16	63	.16	.11	63
PROPDISN brand = N	.09	.46**	64	-.09	.46**	64
PROPDISL brand = L	-.20	.06	62	.15	.12	62
LIKANSUR	.21	.002	188	-.14	.024	188
LIKGRSIZ	.01	.43	188	-.07	.18	188
LIKWATMP	-.11	.06	188	.08	.13	188
LIKSUDS	-.18	.007	188	.13	.04	188
LIKPHOS	-.14	.02	188	.08	.15	188

* All of the correlations presented in this table are between the underlying variables, regardless of the directional coding that was used for those variables.

** Two-tailed P(r) H_0: $\rho = 0$.

Observed relationships between reported elaboration and detergent ratings:

One of the six observed correlations of reported elaboration and detergent ratings was in the non-predicted direction--the correlation of PROPPOS and RATN. However, this correlation was non-significant.

The only significant correlations were those in the predicted direction for PROPNEG and PROPPOS with RATL. Perhaps detergent ratings are less sensitive to changing subject beliefs than are the likelihood of excellence measures. Ss may also have felt compelled to duplicate the ratings as they had been presented in the formula-based test for brands M and N. RATL could be shifted more without approaching the Excellent/ Grade AA category.

Observed relationships between reported elaboration and proportional shift in brand rating:

Reported elaboration was a poor predictor of the proportional shift in a brand's rating. One of the correlations was in the direction opposite to that predicted. However, this correlation was non-significant. One correlation approached significance--that of PROPNEG with PROPDISL. This would indicate that as negative statements increase as a proportion of total statements, the shift in position of the rating of Brand L as a proportion of the distance to the Excellent/Grade AA category would decrease.

Observed Relationship Between Reported Elaboration and Measures of Acceptance:

The pattern of reported elaboration was related to several message-specific and test-specific measures of acceptance.

Both PROPNEG and PROPPOS were related to LIKANSUR and to LIKSUDS. As the perceived likelihood of anionic surfactant content influencing detergent cleaning ability increased, so did PROPNEG. PROPPOS decreased under the same circumstances. As the perceived likelihood of the sudsiness of a detergent determining its cleaning ability increased, PROPNEG decreased and PROPPOS increased.

PROPNEG was also significantly correlated with LIKWATMP and LIKPHOS. As these two likelihoods increased, the proportion of all statements which were negative decreased.

With the exception of LIKGRSIZ, all of the measures related to the content of the test or the content of the message were related to the pattern of reported elaborations. The stronger the belief in message-based factors, the lower the PROPNEG and the higher the PROPPOS. The stronger the belief in one test-based factor (LIKANSUR), the greater the PROPNEG and the lower the PROPPOS. This was as predicted. Unfortunately, as correlational evidence, these findings are compatable with beliefs affecting reported elaboration, with reported elaboration affecting beliefs, or with both the likelihood measures and the reported elaboration measures tapping the underlying cognitive structure. Overall, reported elaboration (especially PROPNEG) was a good predictor of acceptance and vice versa.

The dominance of PROPNEG in predicting structural measures is similar to the results of other researchers (see Table 8). "The underlying assumption of the counterargument model is that counterargument is the primary mediator of acceptance, and therefore that variance in acceptance can be explained satisfactorily solely in terms of the volume of spontaneous counterargument" (Wright 1973a, p. 57).

Internal Analysis

Table 14 presents the correlations between the manipulation check measures and the criterion measures. These correlations express the relationship between the underlying variables regardless of the directionality of the scheme which was used for the variables. These correlations are not presented as support for any causal statements regarding the relationships examined, but rather as a structural analysis to determine how the perceived credibility of the source and the perceived assailability of the background information are related to the post-message cognitive structure as expressed through the criterion measures. However, these perceived values no longer represent conditions to which subjects were randomly assigned, but are themselves reflections of the post-message cognitive structure. This internal analysis can provide insight as to what parts of the cognitive structure appear to be related and aid in the understanding of patterns of response. These insights should be helpful in future efforts to study the underlying process (as opposed to structure) associated with communicative episodes.

Relationship Between Reported Elaboration and Manipulation Check Measures:

The most striking feature of the correlations of PROPNEG and PROPPOS with measures of source and "test" evaluation is the strength of these reported elaboration measures relationship with the subjects' evaluation of the source. All of the correlations of PROPNEG and PROPPOS with the measures of source evaluation are in the predicted direction and are highly significant. Only one of the measures of

Table 14. Correlations of Measures of Communicator Credibility and Information Assailability with Reported Elaboration, Detergent Ratings and Acceptance Measures*

		COMBEL	COMTRUST	COMJUST	COMEXPRT	TESTAPP	TESTREP	TESTRUST
PROPNEG	r	-.43	-.54	-.34	-.31	.16	.10	.09
	p	.00005	.00005	.00005	.00005	.01	.08	.11
	n	187	188	188	188	189	187	188
PROPPOS	r	.26	.34	.44	.35	-.11	-.11	-.11
	p	.0002	.00005	.00005	.00005	.07	.08	.07
	n	187	188	188	188	189	187	188
RATM M	r	.51	.41	.16	.27	.15	.11	-.01
(low dis-	p	.00005	.0006	.10	.018	.25**	.40**	.47
crepancy)	n	60	61	61	61	61	60	61
RATN N	r	.24	.33	.34	.29	.004	.08	.10
(moderate	p	.03	.004	.003	.01	.97**	.54	.43**
discrepancy)	n	63	63	63	63	63	63	63
RATL L	r	.19	.30	.35	.27	-.05	-.06	-.06
(high dis-	p	.08	.009	.003	.02	.36	.32	.33
crepancy)	n	60	60	60	60	60	60	60
PROPDISM M	r	.35	.29	.19	.32	-.15	.15	-.06
(low dis-	p	.0025	.01	.07	.006	.13	.24**	.32
crepancy)	n	61	62	62	62	63	61	62
PROPDISN N	r	.31	.21	.18	.32	-.003	.01	.06
(moderate	p	.006	.05	.08	.004	.49	.94**	.65**
discrepancy)	n	64	64	64	64	64	64	64
PROPDISL L	r	.18	.26	.17	.23	-.008	-.002	.12
(high dis-	p	.08	.02	.10	.04	.48	.49	.37**
crepancy)	n	62	62	62	62	62	62	62
LIKEXM M	r	.65	.63	.30	.35	.22	-.09	-.14
(low dis-	p	.00005	.00005	.008	.003	.09**	.25	.14
crepancy)	n	61	62	62	62	63	61	62
LIKEXN N	r	.23	.18	.24	.21	-.02	-.14	-.08
(moderate	p	.04	.08	.03	.05	.43	.15	.27
discrepancy)	n	63	63	63	63	63	63	63
LIKEXL L	r	.40	.44	.33	.39	-.05	-.06	-.005
(high dis-	p	.0007	.0003	.005	.001	.34	.31	.49
crepancy)	n	61	61	61	61	61	61	61
LIKANSUR	r	-.004	-.08	-.005	.03	.27	.34	.37
	p	.48	.13	.47	.65**	.0001	.00005	.00005
	n	187	188	188	188	188	187	188
LIKGRSIZ	r	.007	.06	-.003	.04	.24	.18	.29
	p	.93**	.42**	.48	.63**	.0005	.007	.00005
	n	187	188	188	188	188	187	188
LIKWATMP	r	.08	.02	-.02	.01	-.07	-.02	-.02
	p	.15	.38	.78**	.42	.19	.38	.40
	n	187	188	188	188	187	186	187
LIKSUDS	r	.28	.23	.17	.17	-.02	.02	-.007
	p	.00005	.0009	.01	.01	.37	.74**	.46
	n	187	188	188	188	188	187	188
LIKPHOS	r	.17	.21	.14	.13	-.04	.12	.07
	p	.01	.002	.03	.04	.29	.11**	.36**
		187	188	188	188	188	187	187

* All of the correlations presented in this table are between the underlying variables regardless of the directional coding that was used for those variables.

** Two-tailed P (r) H_0: $\rho = 0$

"test" evaluation was significantly related to reported elaboration--whether the "test" was seen as appropriate for government use was significantly related to PROPNEG. Although the correlations of the other test evaluation measures with reported elaboration were not significant, all were in the predicted direction.

These findings suggest that, in this study at least, reported elaboration was primarily a source-oriented phenomenon. The proportion of reported statements which was negative or positive was unrelated to whether subjects perceived the "test" as representative or trustworthy. However, the more appropriate for government use subjects perceived the "test" to be the greater was the proportion of their reported elaborative statements which were negative.

It is possible that the relationship between the test and the manufacturer's reply was not drawn clearly enough in this study. Another possibility is that subjects responded on the basis of source-relevant cues which had been developed over a lifetime, rather than to the situation-specific, prior-congnitive-structure variables introduced by the researcher. Whatever the reason, PROPNEG and PROPPOS were unrelated to TESTREP and TESTRUST.

Relationships of Detergent Ratings with Manipulation Check Measures:

Detergent ratings continued to show the same pattern of correlations as the previous measures. None of the ratings were correlated with the measures of "test" evaluation. Yet, all but one of the correlations of detergent ratings (RATM, RATN, RATL) with the measures of source evaluation were significant and all were in the predicted direction.

Relationships of Proportional Shift in Brand Ratings with Manipulation Check Measures:

An interesting difference emerged between the rating correlations and the proportional shift correlations. Although the PROPSHIF measures were also correlated with COMBEL, COMTRUST, and COMEXPRT (only one correlation was not significant and all were in the predicted direction), none of the correlations with COMJUST were significant. COMJUST is certainly the most indirect measure of source credibility--being more of a situation-specific measure. However, recall that the couterpart of COMJUST, for "test" evaluation TESTAPP, was the only measure of "test" evaluaton that was significantly correlated with any of the the criterion measures discussed thus far (it was correlated with PROPNEG).

Again, PROPSHIF appears to be exclusively a source-related response, as have been all but PROPNEG.

Relationships Between Likelihoods of the Detergents' Excellence and the Manipulation Check Measures:

All but one of the correlations of LIKEX with source evaluation measures were significant and all were in the predicted direction. These findings maintain the pattern that has emerged thus far. With the exception of the correlation of PROPNEG with TESTAPP, all of the significant correlations were between source-relevant variables and reported elaboration and detergent evaluation measures. Subjects appear to respond to the communication without regard for the assailability of the information which was attacked in the message. One possibility is that test evaluations are unrelated to the acceptance of the factors being attacked. However, as will be seen in the examination of

correlations involving LIKANSUR and LIKGRSIZ, that is probably not the case. Another possibility is that the link between the message and the background information was not as clear to the subjects as it was to the experimenter. Another possibility (suggested earlier) is that subjects depended upon source-relevant cues developed over their lifetimes to respond to the communications, rather than comparing the message with the structural variables (which included unfamiliar factors) that were introduced in the experimental setting.

Relationships Between Acceptance Measures for "Test"-Relevant Factors and Manipulation Check Measures:

The correlations of LIKANSUR and LIKGRSIZ represent a reversal of the patterns which have been examined so far. Both of these variables were unrelated to measures of source evaluation. However, both were significantly correlated with TESTAPP, TESTREP and TESTRUST in the predicted direction. So, as subjects perceived the "test" as being more appropriate for government use, more representative of actual cleaning ability, or more trustworthy, the ingredients upon which the "test" was based were perceived as being more likely to determine detergent cleaning ability. This is an obvious indication that "test" evaluation was linked with other parts of the cognitive structure, even though it had almost no association with reported elaboration and no association with detergent evaluations.

Relationships Between Acceptance Measures for Communication-Relevant Factors and Manipulation Check Measures:

LIKWATMP was unrelated to either source evaluation or "test" evaluation. This is probably because there was very little variance in the LIKWATMP measure. Apparently water temperature is a very believable

influence on detergent cleaning ability regardless of the perceived credibility of the source or the evaluation of the "test". However, LIKSUDS and LIKPHOS were significantly correlated with all source evaluations measures in the predicted direction. This pattern maintains the cognitive separation that has appeared throughout this internal analysis.

Summary of the Internal Analysis:

The most striking discovery from the internal analysis was the isolation of seemingly related information in the subjects' responses. The evaluation of the assailability of the background information was almost completely unrelated to any responses other than the measures of acceptance of factors upon which the background positions were based. Likewise, the measures of source credibility were unrelated to the evaluation of "test" relevant factors, yet showed a strong pattern of association with reported elaborative responses, detergent evaluations and message relevant factors (except for LIKWATMP). Unfortunately, the correlational nature of an internal analysis precludes any causal statements relative to the observed relationships. However, several suggestions for more clearly testing the relevant relationships through future research in this area may be based on this analysis.

First, the message should clearly attack those aspects of the initial position that are based upon the background information. Second, the validity of the aspects of the initial position which are to be attacked should hinge directly upon the assailability of the background information. Third, acceptance of the message must be contingent upon the rejection of the background information, rather than simply upon the acceptance of a more viable position. These modifications

should make it more difficult to maintain cognitive separation of these areas and should reveal the effects of the predictor variables through the resolution of the conflicting positions.

Further Suggestions For Improved Research of the Relationships Portrayed in the Elaboration Model of Message Processing

In addition to the suggested changes based on the internal analysis, several other changes should improve research in this area. These additional changes focus on the manipulations that were unsuccessful in this study.

The first, and most obvious, suggested change would be the incorporation of a manipulation check to be administered in a pilot study under the guidelines set forth in Chapter 4. This would allow the modification of the manipulations prior to the commitment of time and other resources.

The manipulation of test assailability was completely unsuccessful. One reason may have been that the manipulation of this variable was buried in the middle of other textual material and may have been completely missed by many of the subjects. Of course, another possibility is that even had it been obvious it may not have had the desired result. Nevertheless, a more readily accessible manipulation would be preferred. Manipulation check measures could assess both whether the manipulation had been noticed and whether when noticed, it had the desired impact.

The manipulation of Source Credibility was partially successful. COMEXPRT was affected by the source credibility manipulation. A comparison of the means (high = 2.49, low = 2.95) showed that they were

significantly different ($t_{176} = 2.63$, $p = 0.005$). Even though the difference was highly (statistically) significant, the practical significance of the difference is questionable since even in the low credibility condition the mean response was slightly greater than "Somewhat likely ...that the manufacturer's response is based on the expertise necessary to evaluate detergent cleaning ability" and the two means are both between the same two categories on a six point scale. Likely had a scale value of 2, and somewhat likely had a scale value of 3. (Very likely was 6).

Steps should be taken to address each of the dimensions in the manipulation of a multidimensional variable. Of course, the researcher should strive for practical, as well as statistical significance in the differences achieved on the manipulation check measures.

Other Suggested Research Based on the Elaboration Model of Message Processing:

One set of relationships from the elaboration model that offer an intriguing research opportunity are those involved with the distraction phenomena (recall the extensive discussion in Chapter 3). Any distraction manipulation should follow the criteria outlined in Chapter 3.

Measures of reported elaboration as well as structural measures could be taken following presentation of the stimulus materials.

Overall, many relationships within the elaboration model of information processing remain as viable and important research possibilities.

APPENDIX

Background Reading For Formula-Based

Cleaning Ability Rating System

[High Credibility Version]

We would now like to tell you about the next task. You now know about the proposed U. S. Government cleaning quality rating system. This system would use three categories: Excellent/Grade AA; Very Good/ Grade A; and Good/Grade B. Three categories were chosen because earlier research showed that consumers prefer and can recall the three categories. A detergent is assigned to a category on the basis of a numerical score. This score is the average of several judgements of clothes washed in that detergent. There is a range of scores that fall in each category as you can see here:

10	9	8	7	6	5	4	3	2	1	0
Excellent/ Grade AA			Very Good/ Grade A			Good/ Grade B			Unacceptable	

Detergents are assigned a quality rating according to whichever range their score falls in. The score not only gives the rating but shows how close the brand is to other categories.

There is another rating program that has been proposed which we would like to get your reactions to as detergent consumers. This rating system would provide a preliminary detergent evaluation program for manufacturers. If the predicted rating was not satisfactory to them, they could then revise the formula prior to the washing tests. The estimate would, nevertheless, be made public as is any other government study under the Freedom of Information Act. The news media might, on occasion, publish the results.

FORMULA-BASED RATING SYSTEM

This system uses the same rating categories, but gets the numerical score by another method. This method uses the chemical and physical characteristics of a detergent to estimate the score that would have been assigned by judges if washing tests were used. This test was developed five years ago. Its procedures are well established and have been approved by the Tests and Proceedures Committee of the American Chemical Society.

Some preliminary tests have been run using samples of new, as yet unmarketed, detergents that were supplied voluntarily by the manufacturers for that purpose. The government agency that performed the tests notified the manufacturers that they wanted to get consumer reaction to this rating system. Since the manufacturers had supplied the samples voluntarily, and because the ratings are unofficial, the manufacturers suggested two conditions: (1) that we disguise the brand names so that you could better react to the rating system without being influenced by other information you may have had about the brands; and (2) that they also use this opportunity to get your reactions to a communication of their own regarding the formula-based preliminary tests.

Now to explain a little bit about the procedure of the formula-based rating system. First, manufacturers would provide samples of their new detergents for testing. Then, the physical and chemical characteristics would be determined. Then, a cleaning ability rating would be estimated according to the levels of the chemical and physical characteristics in that detergent.

TECHNICAL SECTION

HOW THE FORMULA-BASED RATING SYSTEM WORKS

The test works this way--because two factors of a detergent seem to indicate the approximate cleaning quality of a detergent, their levels are measured and then the approximate score is calculated. These two factors are grain size and anionic surfactant content (expressed as a percentage of the net weight).

You may have noticed in examining the detergent boxes that anionic surfactants are mentioned on the boxes, however most people are unfamiliar with this term.

ANIONIC SURFACTANTS

Anionic surfactants are, in fact, wetting agents that break the surface tension of the water around particles of dirt to allow them to be flushed away by the water. As the percentage of these surfactants in the formula increases, the cleaning quality of the detergent increases until the percentage reaches about 20%. After that, additional increases do not noticably increase cleaning quality. They do, however, increase cost.

GRAIN SIZE

Powder grain size affects performance through its effect on how quickly and completely the detergent dissolves. A detergent with very large grain size will not dissolve completely during the wash cycle and often leaves a residue on the finished wash. On the other hand, a detergent with grain size that is too small has a tendency to clot together and also not dissolve completely. It is the medium grain size that is the most effective in contributing to cleaning quality.

HOW GRAIN SIZE AND SURFACTANT CONTENT ARE RELATED TO CLEANING ABILITY

It has been found that powder grain sizes appear to be related to cleaning ability in the manner shown on the following page:

GRAIN SIZE AND CLEANING ABILITY

Grain Size	Detracts From Cleaning Ability ←	Neutral ↓	Contributes To Cleaning Ability →
Extra fine....................	(-1)	///	
Fine..........................			///////// (+3)
Medium........................			//////////// (+4)
Coarse........................			///////// (+3)
Extra coarse..................	(-1)	///	

Note that the number of +'s or -'s for any grain size allows an absolute comparison between grain sizes. For example, two +'s is twice as good as one +, three +'s is three times as good as one +, etc.

It has been found that the percentage of anionic surfactants in a detergent formula is related in the following way to cleaning ability: (Here again the number of +'s allows an absolute comparison between percentage levels).

ANIONIC SURFACTANT CONTENT AND CLEANING ABILITY

Percentage Anionic Surfactant	Neutral ↓	Contributes to Cleaning Ability →
1%....................	//////	(+2)
5%....................	/////////	(+3)
10%...................	////////////	(+4)
15%...................	///////////////	(+5)
20% (and above)........	//////////////////	(+6)

Given the powder grain size and the percentage anionic surfactant contained in a detergent formula, a projected cleaning ability rating can be estimated. In effect, this projection combines information from the two preceeding charts.

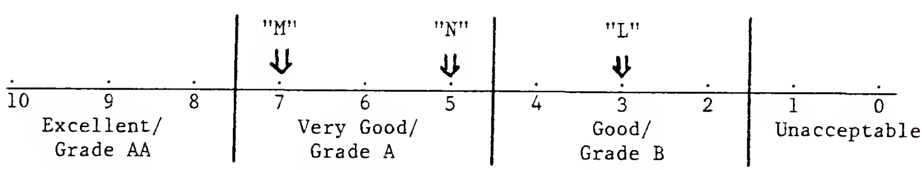
<u>RATING THE NEW BRANDS WITH THE FORMULA-BASED RATING SYSTEM</u>

Below you will see the results of the tests of the samples that were provided by the manufacturers. Because of our agreement, we can only identify the brands by the letters M, N, and L. These brands will be referred to by the same letters for the rest of the study.

The table below shows the grain sizes and surfactant contents for each of these brands along with the estimated scores and ratings that they would receive based on grain size and surfactant content.

Brand	Grain Size	Percentage Anionic Surfactant Content	Estimated Score	Estimated Cleaning Ability Rating
"Brand M"	medium (+4)	5% (+3)	7	Very Good/Grade A
"Brand N"	coarse (+3)	1% (+2)	5	Very Good/Grade A
"Brand L"	fine (-1)	10% (+4)	3	Good/Grade B

It is perhaps easier to visualize these scores by using the line that you saw earlier:



When you have finished reading this section please turn it face down to signal that you have finished. When everyone has finished this section we will hand out a reply from one of the manufacturers. You will be able to refer to this when you read the reply.

Letter from the Manufacturer

[High Credibility Version]

April 2, 1977

Dear Consumer,

We in the Product-Testing Department would like to tell you some of our opinions regarding the formula-based cleaning quality rating system. This system uses grain size and surfactant content to estimate cleaning quality. We agreed to allow the results of our brand on the formula-based test to be used in this study because the outcome is so important to our industry. However, as you will see, we disagree with the results. We have asked that brand names be disguised (ours is called "Brand N") in part because we do not want you to associate what we consider to be erroneous ratings with our brand. In addition, not only are these ratings preliminary, but we think that they are inaccurate and inappropriate.

We are convinced that the formula-based rating system is inaccurate. As Vice-President in charge of Product Testing, I feel confident in assuring you that our new product should receive a rating of Excellent/ Grade AA when the washing tests have been completed. In fact, based on our belief that our product is excellent, we have already gone ahead with the adoption of our present formula and preparations for production.

We also think that the formula-based system is inappropriate. The cleaning ability of a detergent depends on many factors, not just the two used in this test. Some of these other factors are sudsiness, phosphorus content, water softness, and water temperature.

It is because of these points that we don't think that the formula-based test is an adequate measure of detergent cleaning ability. We think that only washing tests are dependable. We are confident that when those results are in, our product will receive a rating of Excellent/Grade AA.

Thank you for your attention and your consideration of our point of view.

Sincerely,

David Richards

David Richards, Ph.D. Chem.
Vice-President
Product-Testing Department

DR/ab
cc: Mr. Gary Sullivan
Mr. P. J. O'Connor

Date____________
Time____________
Sta.#___________

THOUGHT LISTING QUESTIONNAIRE

People tell us that sometimes, when they are reading or listening to a communication, thoughts occur to them that are related to the communication. Other times they may read or hear what was said without really thinking about what was said.

We are now interested in whether or not thoughts related to the communication might have occurred to you while you were reading the manufacturer's reply. On the next page you can record any of these thoughts that you can recall. If you honestly can't recall thinking about what was said, that's fine, simply write NONE on the first line. However, we would like you to write down any thoughts that you can recall having no matter how unimportant or incomplete they may seem to you. Often, even words or phrases that were part of your thoughts can be important. So even if that is all that you can recall, please write them down. Please don't worry about spelling, grammar, or punctuation.

Simply write down the first thought that occurred to you in the first box--the second thought that occurred to you in the second box, etc. Please don't feel that you have to fill a certain number of boxes--we simply want those thoughts that occurred to you as you were reading the manufacturer's reply.

You will have two and one-half minutes to write your thoughts.

PLease turn the page and begin.

Thoughts that occurred to you as you were reading the manufacturer's reply. Please put only one thought per box.

Please go back and read each of the thoughts that you wrote on the previous page. If what you wrote seems to be in favor of what the manufacturer had to say, then write a plus (+) in the left-hand margin next to that thought. If what you wrote seems to be opposed to what the manufacturer had to say, then write a minus (-) in the margin next to that thought. If the thought is neither in favor of nor opposed (is neutral), then you should put a zero (0) in the left-hand margin. Please be sure to rate each thought that you wrote down. Please do not write any additional thoughts at this time.

When you have finished this form, please raise your hand and someone will collect it.

Preference Questionnaire

We would now like you to tell us how you feel about the detergents, M, N, and L that were rated using the formula-based system in considering them for your wash.

On the next page are some comparisons where you can tell how much more you prefer one brand to another for your wash. Here is an example of how to mark your preferences (in this example, we will use imaginary brands):

prefer
Y & Z
equally

prefer Y ← → prefer Z

Y 5 4 3 2 1 0 1 2 3 4 5 Z

If you are indifferent between Y and Z; that is, if you prefer (like) them equally; then you should circle the number "0" in the middle where the arrow indicates "Prefer Y and Z equally".

If you prefer Y to Z, then you should circle a number on the prefer Y side. The number you circle depends on how much you prefer Y to Z. If you like Y a lot more than Z, then you should circle a high number on the prefer Y side. If you prefer Y just a little bit more than Z, you should circle a low number on the prefer Y side. If, on the other hand, you prefer Z to Y, then you should circle a number on the prefer Z side. Again the number you circle depends on how much you prefer Z to Y.

Consider each pair by itself. Do not refer to the other pairs while you are considering any one pair. Remember, judge each pair of detergents according to how much you prefer one over the other for your wash.

prefer L & N equally

prefer L ← → prefer N

L 5 4 3 2 1 0 1 2 3 4 5 N

prefer M & L equally

prefer M ← → prefer L

M 5 4 3 2 1 0 1 2 3 4 5 L

prefer N & M equally

prefer N ← → prefer M

N 5 4 3 2 1 0 1 2 3 4 5 M

FORMULA-BASED RATING SYSTEM QUESTIONNAIRE

1. How likely do you think it is that a cleaning ability rating of Excellent/Grade AA is appropriate for each of the following brands? (Circle a number for each brand):

	very likely	likely	somewhat likely	somewhat unlikely	unlikely	very unlikely
Brand M	1	2	3	4	5	6
Brand N	1	2	3	4	5	6
Brand L	1	2	3	4	5	6

2. You have seen and read a lot of information regarding the Brands M, N, and L. Now we would like your own best judgement of what you think the actual cleaning ability rating for these brands should be. First, think about Brand M. What do you think is the actual cleaning ability rating that this brand would be assigned following the washing tests. Indicate this rating on the line below by drawing an arrow straight down to the score you think it would receive on the washing tests. Now write an "M" above the arrow.

Now think about Brand N and draw an arrow for Brand N. Then write an "N" above this arrow.

Finally, think about Brand L. Draw an arrow for Brand L and write an "L" above this arrow.

10	9	8	7	6	5	4	3	2	1	0
	Excellent/ Grade AA			Very Good/ Grade A			Good/ Grade B		Unacceptable	

Please check to be sure that you have drawn and labeled an arrow for each brand.

(This page has been reduced to meet margin requirements).

3. How believable is the manufacturer's claim that its brand deserves a rating of Excellent/Grade AA?

Very Believable	Believable	Somewhat Believable	Somewhat Unbelievable	Unbelievable	Very Unbelievable
1	2	3	4	5	6

4. How trustworthy do you consider the manufacturer's reply to be?

Very Trustworthy	Trustworthy	Somewhat Trustworthy	Somewhat Untrustworthy	Untrustworthy	Very Untrustworthy
1	2	3	4	5	6

5. How justified is the manufacturer's reply?

Very Justified	Justified	Somewhat Justified	Somewhat Unjustified	Unjustified	Very Unjustified
1	2	3	4	5	6

6. How likely is it that the manufacturer's response is based on the expertise necessary to evaluate detergent cleaning ability?

Very Likely	Likely	Somewhat Likely	Somewhat Unlikely	Unlikely	Very Unlikely
1	2	3	4	5	6

7. How representative of actual cleaning ability do you think ratings based on the formula-based test are?

Very Representative	Representative	Somewhat Representative	Somewhat Unrepresentative	Unrepresentative	Very Unrepresentative
1	2	3	4	5	6

8. How appropriate do you think it is for the government to use the formula-based test to estimate cleaning ability ratings?

Very Appropriate	Appropriate	Somewhat Appropriate	Somewhat Inappropriate	Inappropriate	Very Inappropriate
1	2	3	4	5	6

9. Overall, how trustworthy do you consider the formula-based test to be?

Very Trustworthy	Trustworthy	Somewhat Trustworthy	Somewhat Untrustworthy	Untrustworthy	Very Untrustworthy
1	2	3	4	5	6

Please consider <u>your own beliefs and best judgements</u> in answering the following questions.

(3) What is the likelihood that the percentage anionic surfactant determines detergent cleaning ability?

very likely	likely	somewhat likely	somewhat unlikely	unlikely	very unlikely
1	2	3	4	5	6

(4) What is the likelihood that detergent grain size determines detergent cleaning ability?

very likely	likely	somewhat likely	somewhat unlikely	unlikely	very unlikely
1	2	3	4	5	6

(5) What is the likelihood that water temperature determines detergent cleaning ability?

very likely	likely	somewhat likely	somewhat unlikely	unlikely	very unlikely
1	2	3	4	5	6

(6) What is the likelihood that sudsiness determines detergent cleaning ability?

very likely	likely	somewhat likely	somewhat unlikely	unlikely	very unlikely
1	2	3	4	5	6

(7) What is the likelihood that phosphorus content determines detergent cleaning ability?

very likely	likely	somewhat likely	somewhat unlikely	unlikely	very unlikely
1	2	3	4	5	6

When you have finished this questionnaire please raise your hand and someone will collect it.

Criteria Used by Judges for Coding Protocol Responses into Categories:

Support Arguments (SA)

(1) Declarative statements in favor of the brand's deserving a cleaning ability rating of Excellent/Grade AA.

(2) Statements opposed to the government's formula-based test.

(3) Statements supporting the veracity of a specific argument in the letter from the manufacturer.

Source Enhancement (SE)

(1) Statements affecting the credibility of the manufacturer--expressing trust or confidence in the manufacturer.

(2) Statements expressing approval for the overall approach of the manufacturer in the letter.

Counterarguments (CA)

(1) Declarative statements directed against the brand's deserving an Excellent/Grade AA rating.

(2) Statements supporting the government's formula-based test.

(3) Statements which challenged the accuracy or validity of a specific argument in the letter from the manufacturer.

(4) Statements which suggested an alternative to the manufacturer's proposal.

Source Derogation (SD)

(1) Statements affecting the credibility of the manufacturer--expressing distrust or derogation of the letter or the manufacturer.

(2) Statements expressing dislike for the overall approach of the manufacturer in this letter.

(3) Expressions of ulterior motives of the manufacturer.

Neutral (N)

(1) Statements unrelated to the topic, to the formula-based test, or to the manufacturer.

(2) Simple descriptions of the manufacturer's position.

Debriefing

"The study is really broken into two parts. What we have tried to do in the first part is evaluate a proposed new marketing program in which consumers might be given some additional information about how well each brand performs. You either saw a label on which cleaning ability was rated or, if you were randomly assigned to the control group, you did not see such a rating. As we said earlier, the program is now under evaluation. If it is adopted, the cleaning ability of each brand will be rated and appear on its box.

All brands have not yet been rated, so some of the information you may have seen here represents the best judgement of the program staff. Since a manufacturer can change a product's formula at any time, you might want to continue to use your experience and best judgement in selecting a detergent until such time as final ratings might appear on the boxes.

Industry may respond negatively to this first program. In the second part we wanted to get your reactions to the industry opposition to another rating program. The brands that were used are fictitious. Although we know that anionic surfactant content and grain size are related to cleaning ability, the exact relationship is not known. This was used to establish different ratings for the different brands, since your reactions could depend on the brand rating.

This is a very important study, and we very much appreciate your willingness to provide the essential consumer reactions to the proposed program. We hope it has been an interesting and enjoyable experience for each of you. Now we'd like to ask each of you an important favor.

Other consumers in the Gainesville area have been invited to participate in the study. Some will probably be your neighbors and friends. It is very important that you do not discuss the study with them, since that could affect their behavior and the answers they give us. The Gainesville study will be completed on May 18. So after that date, feel free to talk to whomever you want about the study.

If someone asks you about the study before May 18, how about simply telling them that it was a study to get peoples' reactions to how laundry detergents are sold and some new ideas for stocking shelves in supermarkets. As long as you don't mention the cleaning ratings, or talk about the questions we asked, that's fine.

Now, let's get the finances taken care of! We told you that you would receive a gift detergent, but transporting that much detergent was impossible and we didn't know what brands you would choose. Instead, we are going to give you, in money, an amount equal to the most expensive detergent we used in the study. This, of course, is in addition to the contribution being made to the organization you indicated. We are now going to hand out receipt forms for you to sign. When you have signed them, someone will exchange them for the $1.43. As soon as you receive your money the study is completed and you can go. Again, thank you very much for your help and cooperation."

Table 15. Cell Means for the Likelihoods Associated with Message- and Test-Specific Factors (CC = communicator credibility).

		MESSAGE						NO MESSAGE
		Low Discrepancy		Moderate Discrepancy		High Discrepancy		
	Criterion	high CC	Low CC	high CC	low CC	high CC	low CC	
Low Assail-ability	LIKANSUR	2.36	2.43	2.44	2.13	2.00	2.13	2.11
	LIKGRSIZ	2.07	2.29	2.44	2.07	1.80	2.00	2.21
	LIKWATMP	2.36	2.29	1.94	2.47	2.07	2.40	2.53
	LIKSUDS	3.43	3.14	3.31	2.93	3.07	3.27	3.47
	LIKPHOS	2.64	3.00	2.81	2.87	2.80	2.93	3.37
		n = 14 1	n = 14 2	n = 16 3	n = 15 4	n = 15 5	n = 15 6	n = 19 7
High Assail-ability	LIKANSUR	1.94	2.31	2.13	2.39	2.33	2.18	1.57
	LIKGRSIZ	2.11	2.06	1.73	2.00	2.40	2.18	1.57
	LIKWATMP	1.94	2.00	2.07	2.06	2.27	2.12	2.24
	LIKSUDS	3.39	2.63	3.40	3.22	2.80	3.65	3.52
	LIKPHOS	2.78	2.63	2.67	2.61	2/73	3.24	3.05
		n = 18 8	n = 16 9	n = 15 10	n = 18 11	n = 15 12	n = 17 13	n = 21 14

Table 16. Cell Means for Communicator Credibility Manipulation Check Measures (CC = communicator credibility)

		MESSAGE						NO MESSAGE
		Low Discrepancy		Moderate Discrepancy		High Discrepancy		
		high CC	Low CC	high CC	low CC	high CC	low CC	
	Criterion							
Low Assail-ability	COMBEL	3.14	2.86	3.25	3.13	2.73	4.00	--
	COMTRUST	3.00	3.00	3.38	2.93	3.07	3.73	--
	COMJUST	3.00	2.79	2.94	2.47	2.87	3.40	--
	COMEXPRT	2.71	2.57	3.06	2.67	2.53	3.60	--
		n = 14	n = 14	n = 16	n = 15	n = 15	n = 15	
		1	2	3	4	5	6	7
High Assail-ability	COMBEL	2.65 (17)*	2.75	2.73	3.06	3.07	3.11	--
	COMTRUST	2.72	2.88	2.88	2.80	3.11	3.07	--
	COMJUST	2.67	2.44	2.53	2.94	2.60	3.06	--
	COMEXPRT	2.39	2.69	2.20	3.00	3.07	3.12	--
		n = 18	n = 16	n = 15	n = 18	n = 15	n = 17	
		8	9	10	11	12	13	14

*All means within a cell are based upon the same n except as shown in parentheses.

Table 17. Cell Means for Test Evaluation Measures (CC = communicator credibility).

		MESSAGE						NO MESSAGE
		Low Discrepancy		Moderate Discrepancy		High Discrepancy		
		high CC	Low CC	high CC	low CC	high CC	low CC	
	Criterion							
Low Assail-ability	TESTACC	--	--	--	--	--	--	2.68
	TESTAPP	2.87 (15)*	3.07	3.25	2.53	2.73	3.07	2.74
	TESTREP	2.57	2.50	2.56	2.47	2.93	2.53	2.58
	TESTRUST	2.64	2.71	3.00	2.73	2.93	2.93	2.53
		n = 14	n = 14	n = 16	n = 15	n = 15	n = 15	n = 19
		1	2	3	4	5	6	7
High Assail-ability	TESTACC	--	--	--	--	--	--	2.52
	TESTAPP	2.33	3.06	3.27	2.72	2.27	2.37	2.00
	TESTREP	2.47 (17)*	2.75	3.00	2.61	2.53	2.59	2.43
	TESTRUST	2.39	2.69	3.40	2.78	2.53	2.71	2.43
		n = 18	n = 16	n = 15	n = 18	n = 15	n = 17	n = 21
		8	9	10	11	12	13	14

*All means within a cell are based upon the same n except as shown in parentheses.

LIST OF REFERENCES

Abelson, Robert P. (1976), "Script Processing in Attitude Formation and Decision Making," in John S. Carrol and John W. Payne, Eds., Cognition and Social Behavior, Lawrence Erlbaum Associates, Hillsdale, 33-45.

_________ (1968), "A Summary of Hypotheses on Modes of Resolution," in R. P. Abelson, Elliot Aronson, William J. McGuire, Theodore M. Newcomb, Milton J. Rosenberg and Percy H. Tannenbaum , Eds., Theories of Cognitive Consistency: A Sourcebook, Rand Mc Nally, Chicago, 716-720.

_________ (1959), "Modes of Resolution of Belief Dilemmas," The Journal of Conflict Resolution, 3, 343-352, Reprinted in Joel B. Cohen, Ed., Behavioral Science Foundations of Consumer Behavior, (1972), The Free Press, New York, 275-282.

Adair, John C. (1973), The Human Subject: the Social Psychology of the Psychological Experiment, Little, Brown and Co., Boston.

Ahtola, Olli T. (1975), "The Vector Model of Preferences: An Alternative to the Fishbein Model," Journal of Marketing Research, 12, 52-59.

Aronson, Elliot (1966), "Avoidance of Inter-subject Communication," Psychological Reports, 19, 238.

Atkinson, R. C. and R. M. Shiffrin (1968), "Human Memory: A Proposed System and its Control Processes," in Kenneth W. Spence and Janet Taylor Spence, Eds., The Psychology of Learning and Motivation, Vol. 2, Academic Press, New York, 89-195.

Aune, Bruce (1967), Knowledge, Mind, and Nature, Random House, New York.

Ausubel, David P. (1962), "A Subsumption Theory of Meaningful Verbal Learning and Retention," The Journal of General Psychology, 66, 213-224.

_________ (1963), "Cognitive Structure and the Facilitation of Meaningful Verbal Learning," Journal of Teacher Education, 14 (2), 217-222.

Bakan, David (1954), "A Reconsideration of the Problem of Introspection," Psychological Bulletin, 51 (2), 105-118.

_________ (1955), "The General and the Aggregate: a Methodological Distinction," Perceptual and Motor Skills, 5, 211-212.

Baker, Michael J. and Gilbert A. Churchill, Jr. (1977), "The Impact of Physically Attractive Models on Advertising Evaluations," Journal of Marketing Research, 14, 538-555.

Baron, Robert S. and Norman Miller (1969), "Credibility, Distraction, and Counterargument in a Forewarning Situation," Proceedings, 77th Annual Convention, American Psychological Association, 411-412.

__________, Penny H. Baron, and Norman Miller (1973), "The Relation between Distraction and Persuasion," Psychological Bulletin, 80, 310-323.

Bartlett, Sir Frederic C. (1932), Remembering: A Study in Experimental and Social Psychology, 1964 printing, The Syndics of the Cambridge University Press, Cambridge.

Bawden (1900), "A Study of Lapses," Psychological Review, Monograph Supplement III (4), p. 154 et seq., cited in Pintner (1913).

Begg, Jan and Wayne A. Wilkegren (1974), "Rentention Functions for Syntactic and Lexical vs Semantic Information in Sentence Recognition Memory," Memory & Cognition, 2 (2), 353-359.

Berlyne, D. E. (1968), "The Motivational Significance of Collative Variables and Conflict," in Robert P. Abelson, Elliot Aronson, William J. McGuire, Theordore M. Newcomb, Milton J. Rosenberg, and Percy H. Tannenbaum, Theories of Cognitive Consistency: A Sourcebook, Rand McNally, Chicago, 257-266.

Birnbaum, Michael H. and Barbara A. Mellers (1979), "Stimulus Recognition May Mediate Exposure Effect," Journal of Personality and Social Psychology, 37 (3), 391-394.

Bither, Stewart W. (1972), "Effects of Distraction and Commitment on the Persuasiveness of Television Advertising," Journal of Marketing Research, 9, 1-5.

__________ and Peter L. Wright (1973), "The Self-Confidence-Advertising Response Relationship: A Function of Situational Distraction," Journal of Marketing Research, 10, 146-152.

Blackwell, Roger D., James S. Hensel, and Brian Sternthal (1970), "Pupil Dilation: What Does it Measure," Journal of Advertising Research, 10 (4), 15-18.

Blalock, Hubert M. (1961), Causal Inferences in Nonexperimental Research, University of North Carolina Press, Chapel Hill, cited in Cook (1969).

Bobrow, Daniel G. and Donald A. Norman (1975), "Some Principles of Memory Schemata," in Daniel G. Bobrow and Allan Collins, Eds., Representation and Understanding, Academic Press, New York, 131-149.

Borg, Walter, R. and Meredith Damien Gall (1979), Educational Research: An Introduction, 3rd Edition, Longman, New York.

Boring, Edwin G. (1953), "A History of Introspection," Psychological Bulletin, 50 (3), 169-189.

_________ (1969), "Perspective: Artifact and Control," in Robert Rosenthal and Ralph L. Rosnow, Eds., Artifact in Behavioral Research, Academic Press, New York, 1-11.

Boyd, Harper W., Michael L. Ray, and Edward C. Strong (1972), "An Attitudinal Framework for Advertising Strategy," Journal of Marketing, 36, 27-33.

Bracht, Glenn H. and Gene V. Glass (1968), "The External Validity of Experiments," American Educational Research Journal, 5 (4), 437-474.

Brandt, Lewis W. (1967), "Wanted: A Representative Experiment," Contemporary Psychology, 12 (4), 192-193.

Brehm, J. W. (1966), A Theory of Psychological Reactance, Academic Press, New York, cited in Ralph L. Rosnow and Daniel J. Davies, Eds., (1977), "Demand Characteristics and the Psychological Experiment," Etcetera, 34 (3), 301- 313.

Broadbent, D. E. (1952), "Speaking and Listening Simultaneously," Journal of Experimental Psychology, 43, 267-273.

_________ (1958), Perception and Communication, Pergamon Press, New York.

Brock, Timothy C. (1967), "Communication Discrepancy and Intent to Persuade as Determinants of Counterargument Production," Journal of Experimental Social Psychology, 3, 296-309.

Brockway, John, Donna Chmielewski, and Charles N. Cofer (1974), "Remembering Prose: Productivity and Accuracy Constraints in Recognition Memory," Journal of Verbal Learning and Verbal Behavior, 13, 194-208.

Brodbeck, M. (1966), "Mental and Physical: Identity Versus Sameness," in P. K. Feyerabend and G. Maxwell, Eds., Mind, Matter, and Method: Essays in Philosophy and Science in Honor of Herbert Feigl, University of Minn. Press, Minneapolis, 57-58, cited in Natsoulas 1970.

Brooks, Lee R. (1968), "Spatial and Verbal Components of the Act of Recall," Canadian Journal of Psychology, 22 (5), 349-368.

Bruner, Jerome S. (1957), "On Perceptual Readiness," Psychological Review, 64 (2), 123-152.

Cacioppo, John T., Curt A. Sandman and Barbara B. Walker (1978), "The Effects of Operant Heart Rate Conditioning on Cognitive Elaboration and Attitude Change," Psycholophysiology, 15 (4), 330-338.

_________ and Richard E. Petty (1979), "Effects of Message Repetition and Position on Cognitive Response, Recall, and Persuasion," Journal of Personality and Social Psychology, 37 (1), 97-109.

Calder, Bobby J. (1978), "Cognitive Response, Imagery, and Scripts: What is the Cognitive Basis of Attitude," in H. Keith Hunt, Ed., Advances in Consumer Research, Vol. V, Association for Consumer Research, Ann Arbor, 630-634.

_________, Chester A. Insko, and Ben Yandell (1974), "The Relation of Cognitive and Memorial Processes to Persuasion in a Simulated Jury Trial," Journal of Applied Social Psychology, 4 (1), 62-93.

Campbell, Donald T. (1957), "Factors Relevant to the Validity of Experiments in Social Settings," Psychological Bulletin, 54, 297-312, reprinted in William J. Gephart and Robert B. Ingle (1969), Educational Research: Selected Readings, Charles E. Merrill, Columbus, Ohio, 263-284.

_________ (1960), "Recommendations for APA Test Standards Regarding Construct, Trait, or Discriminant Validity," American Psychologist, 15, 546-553.

_________ (1969), "Prospective: Artifact and Control," in Robert Rosenthal and Ralph L. Rosnow, Eds., Artifact in Behavioral Research, Academic Press, New York, 351-382.

_________ and Donald W. Fiske (1959), "Convergent and Discriminant Validation by the Multitrait-Multimethod Matrix," Psychological Bulletin, 56, 81-105, Reprinted in William A. Mehrens and Robert L. Ebel, Eds., Principles of Educational and Psychological Measurement, 1967, Rand McNally & Co., 273-302.

_________ and Julian C. Stanley (1963), Experimental and Quasi-experimental Designs for Research, Rand McNally & Co., Chicago.

Carlsmith, J. Merril, Phoebe C. Ellsworth and Elliot Aronson (1976), Methods of Research in Social Psychology, Addison-Wesley, Reading.

Cherry, Colin (1966), On Human Communication, 2nd Edition, M.I.T. Press, Cambridge.

Cialdini, Robert B., Alan Levy, C. Peter Herman, Lynn T. Kozlowski, and Richard E. Petty (1976), "Elastic Shifts of Opinion: Determinants of Direction and Durability," Journal of Personality and Social Psychology, 34 (4), 663-672.

Cicourel, Aaron V. (1974), "Interviewing and Memory," in Colin Cherry, Ed., Pragmatic Aspects of Human Communications, D. Reidel Publishing Co., Dordrecht, Holland, 51-82.

Cofer, Charles N. (1976), "An Historical Perspective," in Charles N. Cofer, Ed., The Structure of Human Memory, W. H. Freeman, San Francisco, 1-14.

_________ (1977), "On the Constructive Theory of Memory," in Ina C. Uzgiris and Fredric Weizmann, Eds., The Structuring of Experience, Plenum Press, New York, 319-341.

_________, Donna L. Chmielewski, and John P. Brockway (1976), "Constructive Processes and the Structure of Human Memory," in Charles N. Cofer, Ed., The Structure of Human Memory, W. H. Freeman, San Francisco, 190-203.

Cohen, Joel B. (1974), "Toward an Integrated Use of Expectancy-Value Attitude Models," in G. David Hughes and Michael L. Ray, Eds., Buyer/Consumer Information Processing, University of North Carolina Press, Chapel Hill, 331-346.

_________ (1979), "Exploring Attitude Construct Validity: or Are We?", in William L. Wilkie, Ed., Advances in Consumer Research, Vol VI, Ann Arbor: Assoication for Consumer Research, 303-306.

_________ and Marvin E. Goldberg (1970), "The Dissonance Model in Post-Decision Product Evaluation," Journal of Marketing Research, 7, 315-321.

_________ and Michael J. Houston (1972), "Cognitive Consequences of Brand Loyalty," Journal of Marketing Research, 9, 97-99.

Conrad, Carol (1974), "Context Effects in Sentence Comprehension: A Study of the Subjective Lexicon," Memory & Cognition, 2 (1A), 130-138.

Cook, Thomas D. (1969), "Competence, Counterarguing, and Attitude Change," Journal of Personality, 37, 342-358.

_________ and Donald T. Campbell (1976), "The Design and Conduct of Quasi-Experiments and True Experiments in Field Settings," in Marvin D. Dunnette, Ed., Handbook of Industrial and Organizational Psychology, Rand McNally, Chicago, 223-326.

_________, J. R. Bean, B. J. Calder, R. Frey, M. L. Krovetz, and S. R. Reisman (1970), "Demand Characteristics and Three Conceptions of the Frequently Deceived Subject," Journal of Personality and Social Psychology, 14, 185-194.

Cozby, Paul C. (1977), Methods in Behavioral Research, Mayfield Publishing Co., Palo Alto.

Craik, Fergus I. M. and Robert S. Lockhart (1972), "Levels of Processing: A Framework for Memory Research," Journal of Verbal Learning and Verbal Behavior, 11, 671-684.

__________ and Endel Tulving (1975), "Depth of Processing and the Retention of Words in Episodic Memory," Journal of Experimental Psychology: General, 104 (3), 268-294.

__________ and Michael J. Watkins (1973), "The Role of Rehearsal in Short-Term Memory," Journal of Verbal Learning and Verbal Behavior, 12, 599-607.

Cronbach, Lee J. (1975), "Beyond the Two Disciplines of Scientific Psychology," American Psychologist, 30 (2), 116-127.

__________ and Paul E. Meehl (1955), "Construct Validity in Psychological Tests," Psychological Bulletin, 52 (4), 281-302.

Deutsch, J. A. and D. Deutsch (1963), "Attention: Some Theoretical Considerations," Psychological Review, 70 (1), 80-90.

Dooling, D. James and Robert E. Christiaansen (1977), "Episodic and Semantic Aspects of Memory for Prose," Journal of Experimental Psychology: Human Learning and Memory, 3 (4), 428-436.

Dyer, Jean Royer (1979), Understanding and Evaluating Educational Research, Addison-Wesley, Reading.

Einstein, Albert (1954), A. Einstein, Ideas and Opinions, Crown Publishers, Inc. New York, 1954, cited in Alexsandr Nikolaevick Sokolov, 1972, Inner Speech and Thought, Plenum Press, New York.

Feigl, Herbert (1959), "Philosophical Embarrassments of Psychology," American Psychologist, 14 (3), 115-128.

Festinger, Leon (1953), "Laboratory Experiments, " in Leon Festinger and Daniel Katz, Eds., Research Methods in the Behavioral Sciences, Holt, Rinehart and Winstor, New York, 136-172.

__________ (1954), "A Theory of Social Comparison Processes," Human Relations, VII (2), 117-140.

__________ (1957), A Theory of Cognitive Dissonance, Row, Peterson and Co., Evanston, Ill.

__________ and Nathan Maccoby (1964), "On Resistance to Persuasive Communications," Journal of Abnormal and Social Psychology, 68 (4), 359-366.

Fillenbaum, S. (1966), "Prior Deception and Subsequent Experimental Performance: the 'Faithful' Subject," Journal of Personality and Social Psychology, 4, 532-537, cited in Stephen J. Weber and Thomas

D. Cook (1972), "Subject Effects in Laboratory Research: an Examination of Subject Roles, Demand Characteristics, and Valid Inference," Psychological Bulletin, 77 (4), 273-295.

Fishbein, Martin and Icek Ajzen (1975), Belief, Attitude, Intention and Behavior. An Introduction to Theory and Research, Addison-Wesley Reading.

Franks, Jeffery J. and John D. Bransford (1972), "The Acquisition of Abstract Ideas," Journal of Verbal Learning and Verbal Behavior, 11, 311-315.

Freedman, Jonathan L. and David O. Sears (1965), "Warning, Distraction, and Resistance to Influence," Journal of Personality and Social Psychology, 1 (3), 262-266.

Friedman, Neil (1967), The Social Nature of Psychological Research: the Psychological Experiment as a Social Interaction, Basic Books, New York.

Fromkin, Howard L. and Siegfried Streufert (1976), "Laboratory Experimentation," in Marvin D. Dunnette, Ed., Handbook of Industrial and Organizational Psychology, Rand McNally, Chicago, 415-465.

Fulgosi, A. and J. P. Guilford (1968), "Short-Term Incubation in Divergent Production," American Journal of Psychology, 81, 241-246.

Gardner, David M. (1966), "The Effect of Divided Attention on Attitude Change Induced by a Persuasive Marketing Communication," in Raymond M. Haas, Ed., Science, Technology, and Marketing, American Marketing Association, Chicago, 532-540.

__________ (1970), "The Distraction Hypothesis in Marketing," Journal of Advertising Research, 10 (6), 25-30.

Glucksberg, Sam and George N. Cowen, Jr. (1970), "Memory for Non-attended Auditory Material," Cognitive Psychology, 1, 149-156.

Goldman, William (1973), The Princess Bride, Ballantine Books, New York.

Gorn, Gerald J. and Marvin E. Goldberg (1977), "The Impact of Television Advertising on Children from Low Income Families," Journal of Consumer Research, 4 (2).

Gottsdanker, Robert (1978), Experimenting in Psychology, Prentice-Hall, Inc., Englewood Cliffs.

Greenwald, Anthony G. (1968), "Cognitive Learning, Cognitive Response to Persuasion, and Attitude Change," in Anthony G. Greenwald, Timothy Brock and Thomas Ostrom, Eds., Psychological Foundations of Attitudes, Academic Press, New York, 147-170.

Griggs, Richard A. (1976), "Recognition Memory for Deducible Information," Memory & Cognition, 4 (5), 643-647.

Haaland, Gordon A. and M. Venkatesan (1968), "Resistance to Persuasive Communications: An Examination of the Distraction Hypothesis, Journal of Personality and Social Psychology, 9 (2), 167-170.

Hardyck, Curtis D., Lewis F. Petrinovich, and Delbert W. Ellsworth (1966), "Feedback of Speech Muscle Activity during Silent Reading: Rapid Extinction," Science, 154 (3755), 1467-1468.

Hayes-Roth, Barbara (1977), "Evolution of Cognitive Structures and Processes," Psychological Review, 84 (3), 260-378.

Hess, Eckhard H. (1968), "Pupilometrics," in Frank M. Bass, Charles W. King and Edgar A Pessemier, Eds., Applications of the Sciences in Marketing Management, John Wiley and Sons, Inc., New York.

Hovland, Carl I., Irving L. Janis, and Harold H. Kelley (1953), Communication and Persuasion, Yale University Press, New Haven.

Howe, Michael J. A. (1970), "Repeated Presentation and Recall of Meaningful Prose," Journal of Educational Psychology, 61 (3), 214-219.

Insko, Chester A., William Turnbull, and Ben Yandell (1974), "Facilitative and Inhibiting Effects of Distraction on Attitude Change," Sociometry, 37 (4), 508-528.

Jacobson, Edmund (1973), "Electrophysiology of Mental Activities and Introduction to the Psychological Process of Thinking," in F. J. McGuigan and R. A. Schoonover, Eds., The Psychophysiology of Thinking, Academic Press, New York, 3-31.

Jones, Edward E. and Harold B. Gerard (1967), Foundations of Social Psychology, John Wiley & Sons, Inc., New York.

Jung, John (1971), The Experimenter's Dilemma, Harper & Row, New York.

Kaplan, Abraham (1964), The Conduct of Inquiry: Methodology for Behavioral Science, Chandler Publishing Co., San Francisco.

Katz, Daniel (1972), "Some Final Considerations about Experimentation in Social Psychology," in Charles G. McClintock, Ed., Experimental Social Psychology, Holt, Rinehart and Winston, Inc., New York, 549-561.

Keating, John P. and Timothy C. Brock (1974), "Acceptance of Persuasion and the Inhibition of Counterargumentation Under Various Distraction Tasks," Journal of Experimental Social Psychology, 10, 301-309.

__________ and Bibb Latané (1972), "Distorted Television Reception Distraction, and Attitude Change, " Proceedings, 80th Annual Convention, American Psychological Association.

Keele, Steven W. (1973), Attention and Human Performance, Goodyear Publishing Company, Inc. Pacific Palisades, California.

Kelman, Herbert C. (1953), "Attitude Change as a Function of Response Restriction," Human Relations, Vol. VI (3), 185-214.

__________ (1961), "Processes of Opinion Change," Public Opinion Quarterly, 25, 57-78, Reprinted in Joel B. Cohen, Ed., (1972) Behavioral Science Foundations of Consumer Behavior, The Free Press, New York, 293-305.

__________ (1967), "Human Use of Human Subjects: The Problem of Deception in Social Psychological Experiments," Psychological Bulletin, 67(1), 1-11.

__________ and Reuben M. Baron (1968), "Inconsistency as a Psychological Signal," in Robert P. Abelson, Elliot Aronson, William J. McGuire, Theodore M. Newcomb, Milton J. Rosenberg, and Percy H. Tannenbaum, Eds., Theories of Cognitive Consistency: A Sourcebook, Rand McNally, Chicago, 331-336.

Keppel, Geoffrey (1973), Design and Analysis: A Researcher's Handbook, Prentice-Hall, Inc., Englewood Cliffs.

Kiesler, Charles A., Barry E. Collins, and Norman Miller (1969), Attitude Change, John Wiley & Sons, New York.

Kiesler, Sara B. and Roberta B. Mathog (1968), "Distraction Hypothesis in Attitude Change: Effects of Effectiveness," Psychological Reports, 23, 1123-1133.

Kintsch, Walter (1974), The Representation of Meaning in Memory, Lawrence Erlbaum Associates, Inc., Hillsdale.

__________ (1975), "Memory for Prose," in Charles N. Cofer, Ed., The Structure of Human Memory, W. H. Freeman, San Francisco, 90-113.

__________ (1977), "On Comprehending Stories," in Marcel Adam Just and Patricia A. Carpenter, Eds., Cognitive Processes in Comprehension, Lawrence Erlbaum Associates, Hillsdale, 33-88.

Kintz, B. L., D. J. Delprato, D. R. Mettee, C. E. Persons and R. H. Schappe (1965), "The Experimenter Effect," Psychological Bulletin, 63 (4), 223-232.

Kirk, Rober E. (1968), Experimental Design: Procedures for the Behavioral Sciences, Brooks/Cole, Belmont.

Kolers, Paul A. and David J. Ostry (1974), "Time Course of Loss of Information Regarding Pattern Analyzing Operations," Journal of Verbal Learning and Verbal Behavior, 13, 613-625.

Korman, Abraham K. (1974), The Psychology of Motivation, Prentice-Hall, Englewood Cliffs.

Kruglanski, Arie W. (1973), "Much Ado About the 'Volunteer' Artifacts," Journal of Personality and Social Psychology, 28 (3), 348-354.

_________ (1975a), "The Human Subject in the Psychology Experiment: Fact and Artifact," in Leonard Berkowitz, Ed., Advances in Experimental Social Psychology, Vol. 8, Academic Press, New York, 101-147.

_________ (1975b), "Context, Meaning and the Validity of Results in Psychological Research," British Journal of Psychology, 66 (3), 373-382.

_________ and Zohar Eilam (1974), "Overcoming 'Subject Artifacts': A Critique," British Journal of Social and Clinical Psychology, 13, 9-16.

_________ and Moshe Kroy (1976), "Outcome Validity in Experimental Research: A Re-conceptualization," Representative Research in Social Psychology, 7, 116-178.

Krugman, Herbert E. (1966), "The Measurement of Advertising Involvement," Public Opinion Quarterly, 30, 583-596.

_________ (1971), "Brain Wave Measures of Media Involvement," Journal of Advertising Research, 11 (1), 3-9.

LaBerge, David and S. Jay Samuels (1974), "Toward a Theory of Automatic Information Processing in Reading," Cognitive Psychology, 6, 293-323.

Leavitt, Clark, Charles Waddell, and William Wells (1970), "Improving Day-After Recall Techniques," Journal of Advertising Research, 10 (3), 13-17.

Lindsay, Peter H. and Donald A Norman (1972), Human Information Processing, Academic Press, New York.

_________ and _________ (1977), Human Information Processing, 2nd Edition, Academic Press, New York.

Locke, John (1690), "Essay Concerning Human Understanding," London: Basset, Cited in Edwin G. Boring (1953), "A History of Introspection," Psychological Bulletin, 50 (3), 169-189.

Lund, F. H. (1925), "The Psychology of Belief: IV. The Law of Primacy in Persuasion," Journal of Abnormal Social Psychology, 20, 183-191,

Lutz, Richard J. and John L. Swasy (1977), "Integrating Cognitive Structure and Cognitive Response Approaches to Monitoring Communications Effects," in William D. Perreault, Jr., Ed., Advances in Consumer Research, Vol. IV, Association for Consumer Research, Atlanta, 363-370.

Lyon, Jon G. and Hans-H. Waengler (1976), Covert Articulation in Adult Listeners: Its Relationship to Listening Ability and Graded Verbal Stimuli, Helmut Buske Verlag, Hamburg.

Maccann Haslett, Deirdre (1976), "Distracting Stimuli: Do They Elicit or Inhibit Counterargumentation and Attitude Shift," European Journal of Social Psychology, 6 (1), 81-94.

Maccoby, Nathan and Donald F. Roberts (1972), "Cognitive Processes in Persuasion," in Russell I. Haley, Ed., Attitude Research in Transition, American Marketing Association, 258-278.

MacCorquodale, Kenneth and Paul E. Meehl (1948), "On a Distinction Between Hypothetical Constructs and Intervening Variables," Psychological Review, 55 (2), 95-107.

MacLachlan, James and Priscilla LaBarbera (1978), "Time-Compressed TV Commercials," Journal of Advertising Research, 18 (4), 11-15.

Mahoney, Michael J. (1978), "Experimental Methods and Outcome Evaluation," Journal of Consulting and Clinical Psychology, 46 (4), 660-672.

__________ (1976), Scientist as Subject: The Psychological Imperative, Ballinger Publishing Co, Cambridge.

Makay, Donald G. (1973), "Aspects of the Theory of Comprehension, Memory and Attention," Quarterly Journal of Experimental Psychology, 25, 22-40.

Mayer, Richard E. (1977), Thinking and Problem Solving: An Introduction to Human Cognition and Learning, Scott, Foresman and Co., Glenview, Ill.

McCullough, J. Lee and Thomas M. Ostrom (1974), "Repetition of Highly Similar Messages and Attitude Change," Journal of Applied Psychology, 59 (3), 395-397.

McGuigan, Frank J. (1966), Thinking: Studies of Covert Language Processes, Appleton-Century-Crofts, New York.

__________ (1970), "Covert Oral Behavior During the Silent Performance of Language Tasks," Psychological Bulletin, 74 (5), 309-326.

__________ (1973), "Electrical Measurement of Covert Processes as an Explication of 'Higher Mental Events'," in F. J. McGuigan and R. A Schoonover, Eds., The Psychophysiology of Thinking, Academic Press, New York, 343-385.

__________ (1978), Experimental Psychology: A Methodological Approach, 3rd. Ed., Prentice-Hall, Englewood Cliffs.

__________, Barbara Keller, and Eleanor Stanton (1964), "Covert Language Responses During Silent Reading," Journal of Educational Psychology, 55 (6), 339-343.

McGuire, William J. (1966), "Attitudes and Opinions," Annual Review of Psychology, 17, 475-514.

__________ (1968), "Theory of the Structure of Human Thought," in Theories of Cognitive Consistency: A Sourcebook, Robert P. Abelson, Elliot Aronson, William J. McGuire, Theodore M. Newcomb, Milton J. Rosenberg and Percy H. Tannenbaum, Rand McNally, Chicago, 140-162.

Meehl, Paul E. (1967), "Theory-testing in Psychology and Physics: a Methodological Paradox," Philosophy of Science, 34, 103-115.

Milgram, Stanley (1963), "Behavioral Study of Obedience," Journal of Abnormal and Social Psychology, 67, 371-378, cited in James T. Tedeschi and Svenn Lindskold, 1976, Social Psychology. Interdependence, Interaction, and Influence, John Wiley & Sons, New York.

Miller, Norman and Robert S. Baron (1973), "On Measuring Counterarguing," Journal for the Theory of Social Behavior, 3, 101-118.

__________ and Burton H. Levy (1967), "Defaming and Agreeing with the Communicator as a function of Emotional Arousal, Communication Extremity, and Evaluative Set," Sociometry, 30 (2), 158-175.

Mistler-Lachman, Janet L. (1974), "Depth of Comprehension and Sentence Memory," Journal of Verbal Learning and Verbal Behavior, 13, 98-106.

Morton, John (1969), "Interaction of Information in Word Recognition," Psychological Review, 76 (2), 165-178.

Murray, H. G. and J. P. Denny (1969), "Interaction of Ability Level and Interpolated Activity (Opportunity for Incubation) in Human Problem Solving," Psychological Reports, 24, 271-276.

Natsoulas, Thomas (1967), "What are Perceptual Reports About," Psychological Bulletin, 67 (4), 249-272.

__________ (1970), "Concerning Introspective Knowledge," Psychological Bulletin, 73 (2), 89-111.

Neisser, Ulric (1967), Cognitive Psychology, Prentice-Hall, Englewood Cliffs.

Nisbett, Richard E. and Timothy DeCamp Wilson (1977), "Telling More Than We Can Know: Verbal Reports on Mental Processes," Psychological Review, 84 (3), 231-259.

Norman, Donald A. (1976), Memory and Attention, 2nd Edition, John Wiley & Sons, New York.

__________ and Daniel G. Bobrow (1975), "On Data-Limited and Resource-Limited Processes," Cognitive Psychology, 7, 44-64.

__________ and __________ (1976), "On the Role of Active Memory Processes in Perception and Cognition," in Charles N. Cofer, Ed., The Structure of Human Memory, W. H. Freeman, San Francisco, 114-132.

Orne, Martin T. (1961), "On the Social Psychology of the Psychological Experiment: with Particular Reference to Demand Characteristics and their Implications," paper presented at the Symposium, "On the Social Psychology of the Psychological Experiment," American Psychological Assn. Convention, New York, reprinted in Carl W. Backman and Paul F. Second, Eds., (1966), Problems in Social Psychology: Selected Readings, McGraw Hill, New York.

__________ (1969), "Demand Characteristics and the Concept of Quasi-Controls," in Robert Rosenthal and Ralph L. Rosnow, Eds., Artifacts in Behavioral Research, Academic Press, New York, 143-179.

Osgood, Charles E., George J. Suci, and Percy H. Tannenbaum (1957), The Measurement of Meaning, University of Illinois Press, Urbana.

Osterhouse, Robert A. and Timothy C. Brock (1970), "Distraction Increases Yielding to Propaganda by Inhibiting Counterarguing," Journal of Personality and Social Psychology, 15 (4), 344-358.

Payne, Don (1977), "The Biases of Inexact Hypothesis Testing," Etcetera, 34 (3), 330-338.

Peterson, Robert A. (1977), "Consumer Perceptions as a Function of Product Color, Price and Nutrition Labeling," in William D. Perreault, Jr., Ed., Advances in Consumer Research, Association for Consumer Research, Atlanta, 4, 61-63.

Petty, Richard E. (1977), "The Importance of Cognitive Responses in Persuasion," in William D. Perreault, Jr., Ed., Advances in Consumer Research, Association for Consumer Research, Atlanta, 5, 357-362.

__________ and John T. Cacioppo (1977), "Forewarning, Cognitive Responding and Resistance to Persuasion," Journal of Personality and Social Psychology, 35 (9), 645-655.

__________ and __________ (1979), "Issue Involvement Can Increase or Decrease Persuasion by Enhancing Message-Relevant Cognitive Responses," Journal of Personality and Social Psychology, 37 (10), 1915-1926.

__________, Gary L. Wells, and Timothy C. Brock (1976), "Distraction Can Enhance or Reduce Yielding to Propaganda: Thought Disruption Versus Effort Justification," Journal of Personality and Social Psychology, 34 (5), 874-884.

Pintner, Rudolf (1913), "Inner Speech During Silent Reading," Psychological Review, 20, 129-153.

Pitkin, Walter B. (1932), A Short Introduction to the History of Human Stupidity, Simon and Schuster, New York.

Posner, Michael I. (1973), Cognition: An Introduction, Scott, Foresman and Company, Glenview, Illinois.

__________ and Stephen J. Boies (1971), "Components of Attention," Psychological Review, 78 (5), 391-408.

__________ and Raymond M. Klein (1973), "On the Functions of Consciousness," in Sylval Kronblum, Ed., Attention and Performance IV, Academic Press, New York, 21-35.

__________ and Robert E. Warren (1972), "Traces, Concepts, and Conscious Constructions," in Arthur W. Melton and Edwin Martin, Eds., Coding Processes in Human Memory, V. H. Winston & Sons, Washington, 25-43.

Postman, Leo (1975), "Verbal Learning and Memory," Annual Review of Psychology, 26, 291-335.

Pylyshyn, Zenon W. (1973), "What the Mind's Eye Tells the Mind's Brain: A Critique of Mental Imagery," Psychological Bulletin, 80 (1), 1-24.

Reitman, Judith Spencer (1971), "Mechanisms of Forgetting in Short-Term Memory," Cognitive Psychology, 2, 185-195.

__________ (1974), "Without Surreptitious Rehearsal, Information in Short-Term Memory Decays," Journal of Verbal Learning and Verbal Behavior, 13, 365-377.

Richey, Harold W. (1976), "Avoidable Failures of Experimental Procedure," Journal of Experimental Education, 45 (2), 10-13.

Riecken, Henry W. (1962), "A Program for Research on Experiments in Social Psychology," in N. F. Washburne, Ed., Decisions, Values and Groups, Vol. 2, New York, Pergamon Press, reprinted in John Jung (1971), The Experimenter's Dilemma, Harper & Row, New York.

Roberts, Donald F. and Nathan Maccoby (1973), "Information Processing and Persuasion: Counterarguing Behavior," in Sage Annual Review of Communication Research, New Models for Mass Communication Research, Vol. 2, Sage Publications, Beverly Hills, 269-307.

Rokeach, Milton (1968a), Beliefs, Attitudes, and Values: A Theory of Organization and Change, Jossey-Bass, Inc., San Francisco.

__________ (1968b), "A Theory of Organization and Change Within Value-Attitude Systems," Journal of Social Issues, 24 (1), 13-33.

__________ (1968c), "The Role of Values in Public Opinion Research," Public Opinion Quarterly, 32, 547-559.

Rosenberg, Milton J. (1956), "Cognitive Structure and Attitudinal Affect," Journal of Abnormal and Social Psychology, 53, 367-372.

_________ (1969), "The Conditions and Consequences of Evaluation Apprehension," in Robert Rosenthal and Ralph L. Rosnow, Eds., Artifact in Behavioral Research, Academic Press, New York, 279-349.

Rosenblatt, Paul C. (1966), "Persuasion as a Function of Varying Amounts of Distraction," Psychonomic Science, 5 (2), 85-86.

Rosenthal, Robert (1966), Experimenter Effects in Behavioral Research, Appleton-Century-Crofts.

_________ (1977), "Biasing Effects of Experimenters," Etcetera, 34 (3), 253-264.

_________ and Ralph L. Rosnow (1975), The Volunteer Subject, John Wiley & Sons, New York.

Rosnow, Ralph L. and Daniel J. Davis (1977), "Demand Characteristics and the Psychological Experiment," Etcetera, 34 (3), 301-313.

Rule, Brendan Gail and David Rehill (1970), "Distraction and Self-Esteem Effects on Attitude Change," Journal of Personality and Social Psychology, 15 (4), 359-365.

Rumelhart, David E. (1975), "Notes on a Schema for Stories," in Daniel B. Bobrow and Allan Collins, Eds., Representation and Understanding, Academic Press, New York, 211-236.

SAS Institute (1979), SAS User's Guide, 1979 Edition, SAS Institute, Inc., Raleigh.

Schlenker, Barry R. (1974), "Social Psychology and Science," Journal of Personality and Social Psychology, 29 (1), 1-15.

Schneider, Walter and Richard M. Shiffrin (1977), "Controlled and Automatic Human Information Processing: I. Detection, Search, and Attention," Psychological Review, 84 (1), 1-66.

Sears, D. O. and J. L. Freedman (1972), "Selective Exposure to Information: A Critical Review," in J. B. Cohen, Ed., Behavioral Science Foundations of Consumer Behavior, The Free Press, New York, 38-50.

Shank, Roger C. and Robert P. Abelson (1975), "Scripts, Plans, and Knowledge," in Advance Papers of the Fourth International Joint Conference on Artificial Intelligence, Sponsored by the International Joint Council on Artificial Intelligence, Vol. I, Tbilisi, Georgia, USSR, 3-8 September, 151-157.

Sherif, Muzafer and Carl I. Hovland (1961), Social Judgement: Assimilation and Contrast Effects in Communication and Attitude Change, Yale University Press, New Haven.

Shiffrin, Richard M. and Walter Schneider (1977), "Controlled and Automatic Human Information Processing: II. Perceptual Learning, Automatic Attending, and a General Theory," Psychological Review, 84 (2), 127-190.

Silveira, J. (1971), "The Effect of Interruption Timing and Length on Problem Solution and Quality of Problem Processing," Unpublished doctoral dissertation, University of Oregon, cited in Michael I. Posner, 1973, Cognition: An Introduction, Scott, Foresman and Co., Clenview, Ill.

Silverman, Irwin (1977), The Human Subject in the Psychological Laboratory, Pergamon Press, New York.

__________ and C. Robert Regula (1968), "Evaluation Apprehension, Demand Characteristics, and the Effects of Distraction on Persuasibility," The Journal of Social Psychology, 75, 273-281.

Slobin, Dan I. (1974), Psycholinguistics, Scott, Foresman and Company, Glenview, Illinois.

Sokolov, Aleksandr Nikolaevick (1969), "Studies of the Speech Mechanisms of Thinking," in Michael Cole and Irving Maltzman, Eds., A Handbook of Contempory Soviet Psychology, Basic Books Inc., New York, 531-573.

__________ (1972), Inner Speech and Thought, Translated by George T. Onischenko, Plenum Press, New York.

Spelke, Elizabeth, William Hirst, and Ulric Neisser (1976), "Skills of Divided Attention," Cognition, 4, 215-230.

Stricker (1880), "Studien uber die Sprachvorstellungen," Wien, cited in Agnes M. Thorson (1925), "The Relation of Tongue Movements of Internal Speech," Journal of Experimental Psychology, 8 (1), 1-32.

Stricker, Lawrence, Samuel Messick and Douglas N. Jackson (1969), "Evaluating Deception in Psychological Research," Psychological Bulletin, 71 (5), 343-351.

Tedeschi, James T. and Svenn Lindskold (1976), Social Psychology. Interdependence, Interaction, and Influence, John Wiley & Sons, New York.

__________, Barry R. Schlenker, and Thomas V. Bonoma (1971), "Cognitive dissonance: Private Ratiocination or Public Spectacle," American Psychologist, 26, 685-695.

Thorsen, Agnes M. (1925), "The Relation of Tongue Movements to Internal Speech," Journal of Experimental Psychology, 8 (1), 1-32.

Triesman, Anne and Julia Tuxworth (1974), "Immediate and Delayed Recall of Sentences After Perceptual Processing at Different Levels," Journal of Verbal Learning and Verbal Behavior, 13, 38-44.

Tulving, Endel (1972), "Episodic and Semantic Memory," in Endel Tulving and Wayne Donaldson, Eds., Organization of Memory, Academic Press, New York, 381-403.

__________ and Stephen A. Madigan (1970), "Memory and Verbal Learning," Annual Review of Psychology, 21, 437-484.

Underwood, Benton J. and John J. Shaughnessy (1975), Experimentation in Psychology, John Wiley & Sons, Inc., New York.

Vohs, John L. (1964), "An Empirical Approach to the Concept of Attention," Speech Monographs, 31 (3), 355-360.

__________ and Roger L. Garrett (1968), "Resistance to Persuasion: An Integrative Framework," Public Opinion Quarterly, 32 (3), 445-452.

Ward, Scott (1974), "A Discussion of Wright's Paper on Direct Monitoring," in G. D. Hughes and M. L. Ray, Eds., Buyer/Consumer Information Processing, University of North Carolina Press, Chapel Hill.

Watson, John B. (1913), "Psychology as the Behaviorist Views It," Psychological Review, 20, 158-177.

__________ (1920), Is Thinking Merely the Action of Language Mechanisms? (V.)," The British Journal of Psychology, Vol. XI, part 1.

__________ (1966), "What Is Thinking," in F. J. McGuigan, Ed., Thinking: Studies of Covert Language Processes, Appelton-Century-Crofts, New York, 9-16.

Webb, Eugene J., Donald T. Campbell, Richard D. Schwarz and Lee Sechrest (1966), Unobtrusive Measures: Nonreactive Research in the Social Sciences, Rand McNally College Publishing Co., Chicago.

Weber, Stephen J. and Thomas D. Cook (1972), "Subject Effects in Laboratory Research: An Examination of Subject Roles, Demand Characteristics, and Valid Inference," Psychological Bulletin, 77 (4), 273-295.

Williams, G. W. (1963), "Highway Hypnosis: An Hypothesis," International Journal of Clinical and Experimental Hypnosis, 11, 143-151, cited in Thomas Natsoulas (1970), "Concerning Introspective Knowledge," Psychological Bulletin, 73 (2), 89-111.

Wright, Peter L. (1973a), "The Cognitive Processes Mediating Acceptance of Advertising," Journal of Marketing Research, 10, 53-62.

__________ (1973b), "Use of Consumer Judgement Models in Promotion Planning," Journal of Marketing, 37, 27-33.

_________ (1974a), "On the Direct Monitoring of Cognitive Response to Advertising," in G. D. Hughes and M. L. Ray, Eds., Buyer/Consumer Information Processing, University of North Carolina Press, Chapel Hill, 220-248.

_________ (1974b), "Analyzing Media Effects on Advertising Responses," Public Opinion Quarterly, 38, 192-205.

_________ (1975), "Factors Affecting Cognitive Resistance to Advertising," Journal of Consumer Research, 2, 1-9.

Wyer, Robert S., Jr. (1974), Cognitive Organization and Change: An Information Processing Approach, Lawrence Erlbaum Associates, Inc., Potomac.

Zimbardo, Philip, Mark Snyder, James Thomas, Alice Gold, and Sharon Gurwitz (1970), "Modifying the Impact of Persuasive Communications with External Distraction," Journal of Personality and Social Psychology, 16 (4), 669-680.

BIOGRAPHICAL SKETCH

John Vann was born in Melbourne, Florida, in 1944. He majored in Physics at Emory University, receiving his BA degree in 1966. Two years were spent teaching English as a foreign language for the Peace Corps in Afghanistan (1966-1968). John attended Michigan State University where he received his MBA in 1969. After teaching physics and physical science in a Michigan high school, he taught in the Management Science Department at Florida Institute of Technology in Melbourne, Florida for four years before beginning the Ph.D. program at the University of Florida.

I certify that I have read this study and that in my opinion it conforms to acceptable standards of scholarly presentation and is fully adequate, in scope and quality, as a dissertation for the degree of Doctor of Philosophy.

Joel B. Cohen, Chairman
Professor of Marketing

I certify that I have read this study and that in my opinion it conforms to acceptable standards of scholarly presentation and is fully adequate, in scope and quality, as a dissertation for the degree of Doctor of Philosophy.

William L. Wilkie
Professor of Marketing

I certify that I have read this study and that in my opinion it conforms to acceptable standards of scholarly presentation and is fully adequate, in scope and quality, as a dissertation for the degree of Doctor of Philosophy.

Barry R. Schlenker
Associate Professor of Psychology

This dissertation was submitted to the Graduate Faculty of the Department of Marketing in the College of Business Administration and to the Graduate Council, and was accepted as a partial fulfillment of the requirements for the degree of Doctor of Philosophy.

August, 1980

Dean, Graduate School